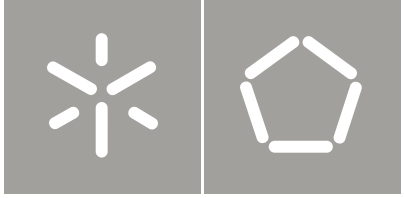
Universidade do Minho
Escola de Engenharia

Nelson Ricardo Lima da Silva

Método de Implementação de SIEMs: Resultados de Experiências Práticas

Método de Implementação de SIEMs: Resultados de Experiências Práticas
Nelson Ricardo Lima da Silva
UMinho | 2011

Outubro de 2011



Universidade do Minho
Escola de Engenharia

Nelson Ricardo Lima da Silva

Método de Implementação de SIEMs: Resultados de Experiências Práticas

Tese de Mestrado
Engenharia e Gestão de Sistemas de Informação

Trabalho efectuado sob a orientação do
Professor Doutor Henrique Dinis dos Santos
Universidade do Minho

Sérgio V. C. Sá
Unisys Portugal

Outubro de 2011

Agradecimentos

Aos meus pais José e Lúcia, pelo amor e carinho que nunca faltou, pelo incentivo incansável, pela forma como me educaram, pelo sentido de responsabilidade que me incutiram e pelos sacrifícios que fizeram para que esta etapa fosse possível...

Ao meu irmão Rafael, pela presença assídua que sempre teve para me acompanhar nos momentos em que a distracção era importante e pela importância que teve ao longo do meu crescimento...

À minha avó Erminda, pela constante preocupação demonstrada...

Ao meu orientador Henrique Santos, pelo seu contributo e principalmente pela forma como me fez encarar todo este trabalho...

Ao meu orientador na Unisys Portugal, pelas oportunidades de aprendizagem e pela confiança que sempre demonstrou nas minhas capacidades...

À Unisys Portugal e em particular à equipa do GOIS pela forma como me receberam e me ajudaram, pela amizade que continuam a demonstrar e pela forma como me ajudaram a evoluir profissionalmente...

Ao João Nunes, ao João Silva e ao David Matslofva, três especialistas em ArcSight SIEM, pelos conhecimentos que me transmitiram e pela paciência em me aturar ao longo dos projectos em que trabalhos juntos...

Ao Alexandre e Conceição, pela forma como me acolheram em Lisboa, assim como aos seus filhos Raquel e Luis, pelos tempos juntos que passamos como irmãos...

À minha namorada Natália, pelo amor demonstrado, pela companhia ao longo dos últimos anos desde a licenciatura, pelos momentos felizes que passamos juntos, pelo acompanhamento que teve à distância no decorrer deste trabalho...

A todas a pessoas que de alguma forma prestaram o seu contributo das mais variadas formas e cujos nomes não foram mencionados anteriormente...

E por fim, a todos aqueles que nunca acreditaram no meu sucesso, a todos aqueles que sempre desejaram que as coisas acabassem mal e a todos aqueles que de alguma forma me tentaram fazer desistir, pela forma como fizeram crescer a minha vontade de os contrariar e mostrar que eu era capaz...

Método de Implementação de SIEMs: Resultados de Experiências Práticas

Resumo

A crescente utilização das Tecnologias de Informação e da Comunicação por parte das empresas, aliada à maior complexidade das relações que estabelecem com os seus parceiros de negócio, expõe cada vez mais os sistemas informáticos a ataques internos e externos. Desta forma surge cada vez mais a necessidade de protecção daqueles sistemas. Os SIEM (Security Information and Event Management) são sistemas cada vez mais usados para proteger a informação que as empresas gerem, seguindo uma estratégia de análise centralizada de múltiplos eventos de segurança, gerados por diversos componentes de segurança.

Devido ao custo elevado dos SIEM bem como à complexidade da sua implementação, torna-se necessário torná-la eficaz, tanto do ponto de vista do tempo de execução das actividades de implementação como da eficiência dos mecanismos de detecção de ataques que estes sistemas permitem. Assim, o principal objectivo deste trabalho é propor um novo método de implementação de SIEM que consolide as melhores práticas e integre a experiência prática do aluno.

O trabalho realizado tem como principais resultados a apresentação de informação detalhada sobre os sistemas de detecção de intrusões e de um método para orientação das actividades de implementação de sistemas de SIEM.

Depois de validado e optimizado, o método proposto neste trabalho será particularmente útil para os profissionais que pretendam entender o processo de implementação de um SIEM bem como obter recomendações práticas para as actividades que necessitará executar de forma a assegurar a implementação adequada destes sistemas.

Implementation Method for SIEMs: Results from Practical Experience

Abstract

The increasing use of Information Technologies and Communications by companies, coupled with the increased complexity of the business relationships that they establish with partners, increasingly exposes their information systems to internal and external attacks. Thus it is increasingly felt the need to protect those systems.

The SIEM systems are increasingly used to protect the information that companies manage. Due to the high cost of SIEM and the complexity of its implementation, it becomes necessary to make implementation effective, both in terms of execution time of implementation activities and in terms of the effectiveness of attack detection mechanisms that these systems allow. Therefore, the main objective of this project is to advance a new method for the implementation of SIEM that consolidates the identified best practices and integrates the work practice of the student.

This dissertation work's main results include the presenting of detailed information on intrusion detection systems and a method to guide the activities of implementing SIEM systems. Once validated and refined, the method proposed in this document will be particularly useful for professionals who wish to understand the process of implementing a SIEM as well as to get practical recommendations for activities that they need to perform to ensure proper implementation.

Índice

Índice de Figuras

Índice de Tabelas

Lista de Siglas e Acrónimos

2FA	*Two Factor Authentication*
APT	*Advanced Persistent Threat*
ASCII	*American Standard Code for Information Interchange*
DDos	*Distributed Denial of Service*
DNS	*Domain Name System*
DoS	*Denial of Service*
EPS	*Events per Second*
HIDS	*Host-Based Intrusion Detection System*
IATAC	*Information Assurance Technology Analysis Center*
IDS	*Intrusion Detection System*
IPS	*Intrusion Prevention System*
ISO	*International Organization for Standardization*
IT	*Information Technology*
NAT	*Network Address Translation*
NIDS	*Network Intrusion Detection System*
NIST	*National Institute of Standards and Technologies*
NTP	*Network Time Protocol*
OPSEC	*Operations Security*
OSI	*Open Systems Interconnection*
PCI DSS	*Payment Card Industry Data Security Standards*
RADIUS	*Remote Authentication Dial In User Service*
RAID	*Redundant Array of Independent Disks*
RAT	*Remote Administration Toolkit*
SEM	*Security Event Management*
SIEM	*Security Information and Event management*
SIM	*Security Information System*
SMTP	*Simple Mail Transfer Protocol*
SSE-CMM	*Systems Security Engineering Capability Maturity Model*
TCP	*Transmission Control Protocol*
VPN	*Virtual Private Network*

Capítulo I - Introdução

Motivação

A crescente utilização das Tecnologias de Informação e da Comunicação por parte das empresas, aliada à maior complexidade das relações que estabelecem com os seus parceiros de negócio, expõe cada vez mais os sistemas informáticos a ataques internos e externos.

Cada vez mais têm vindo a surgir relatos de ataques aos sistemas de empresas conhecidas, com danos de dimensão significativa e que chamam a atenção para a necessidade do investimento em sistemas de detecção e monitorização de intrusões. A seguir apresenta-se a descrição de alguns casos recentes.

RSA Security

Data em que o incidente foi reportado: 17de Março de 2011

"In a March 17 open letter to its customers, RSA executive chairman Art Coviello outlined the preliminary details of what it determined to be "a sophisticated advanced persistent threat" attack that compromised information related to the company's SecurID two-factor authentication (2FA) product."(Amorosi, 2011).

Impacto dos danos: cada *token* da RSA Security tem uma chave de registo associado. Uma vez que esta informação foi roubada é possível clonar qualquer *token.* Tendo em conta que cerca de 30000 organizações em todo o mundo utilizam estes produtos da RSA, a sua credibilidade está neste momento em causa.

Segundo Tankard a RSA Security foi alvo de ataques que tem sido frequente nos últimos anos, categorizados como *Advanced Persistent Threat* (APT). Este ataque, para além de exploração de vulnerabilidades e utilização de *Phishing* direccionado, envolveu ferramentas designadas de *Remote Access Toolkit* (RAT) que foram utilizadas para estabelecer um canal de comunicação entre a rede interna da RSA Security e um servidor remoto gerido pelos atacantes. Estas ferramentas RAT permitem, entre outras coisas, capturar informação de câmaras, equipamentos de rede, pesquisar e gerir ficheiros nos sistemas, gerir as entradas do registo do sistema. Uma vez que estas ferramentas não foram devidamente identificadas na infra-estrutura interna da organização, a informação foi recolhida e transmitida para o tal servidor do atacante que ficou assim na posse da mesma (Tankard, 2011).

De acordo com um dos chefes da RSA, Uri Rivner, o ataque começou por utilizar *Phishing* direccionado aproveitando uma vulnerabilidade do Adobe Flash quando este foi disponibilizado a partir de um ficheiro de *Excel* malicioso. Uma série de emails foram enviados para dois grupos de colaboradores dentro da RSA com o anexo infectado; alguns destes inconscientemente abriram o anexo infectando assim as máquinas e instalando *software* malicioso na infra-estrutura interna da organização. Apesar da utilização de firewalls, sistemas de prevenção de intrusões e outros mecanismos de segurança, estas medidas não foram suficientes para evitar a perda de informação da RSA Security (Amorosi, 2011).

O ataque à RSA Security pode não ser considerado grande em termos de registos efectivamente comprometidos, mas o ataque teve uma repercussão alargada, não só na empresa de segurança mas em todas as empresas e agência governamentais que dependem da tecnologia de autenticação de duplo factor "*SecureID*" para a sua segurança (Rashid, 2011).

Sony

Data em que o incidente foi reportado: 26 de Abril de 2011.

Impacto dos danos: mais de 100 milhões de contas de utilizador roubadas. A Sony estima que este ataque vai ter um custo de cerca de 178 milhões de euros apenas no ano fiscal de 2011

Tipo de dados: nome, morada, endereço de e-mail, credenciais de login e alguma informação de cartões de crédito.

Segundo afirma Rashid, os três serviços de *Cloud* da Sony para os jogos PlayStation, música e vídeo, e jogos online foram comprometidos por atacantes enquanto a companhia estava distraída com um ataque de negação de serviço (DDoS) distribuído de diferentes fontes (Rashid, 2011).

De acordo com Tankard, os ataques à Sony foram realizados durante vários meses tendo sido exploradas uma série de vulnerabilidades, sendo que em muitos dos casos estas vulnerabilidades estavam associadas a *software* desactualizado (Tankard, 2011).

Apesar de os detalhes serem ainda reduzidos dada a continuidade da investigação, a Sony deu a conhecer que entre 17 e 19 de Abril os seus serviços de PlayStation Network and Qriocity foram comprometidos por uma intrusão ilegal e não autorizada na sua rede. Esta intrusão levou ao roubo de milhões de contas de utilizador e informação relacionada com cartões de crédito que estava a ser vendida clandestinamente (Amorosi, 2011).

Operação Aurora

Data em que o incidente foi reportado: 12 de Janeiro de 2010.

A operação Aurora foi um ataque que começou em meados de 2009 e continuou até ao final do mesmo ano. Apenas em Janeiro de 2010, a Google divulgou publicamente este ataque com uma publicação num blogue afirmando que a origem do ataque terá sido na China e teria sido utilizado um ataque do tipo *Advanced Persistent Threat*. Para além da Google também empresas como a Juniper, Symantec, Adobe Systems, entre outras foram vítimas do mesmo ataque. Nesse mesmo blogue Google afirma que duas contas de correio electrónico do *Gmail* foram roubadas e os seus conteúdos foram copiados.

A operação Aurora começou com a exploração de vulnerabilidades "*zero-day*", ou seja, foram exploradas vulnerabilidades que ainda não eram conhecidas e portanto não teriam ainda actualizações de correcção. Estas vulnerabilidades foram identificadas no Internet Explorer assim como outras vulnerabilidades que foram utilizadas no sistema Perforce (Kim Zetter, 2011). Para explorar estas vulnerabilidades foram enviados emails de *Phishing* para colaboradores com hiperligações para determinados sites que continham código malicioso de *JavaScript*, sendo este instalado nas máquinas dos colaboradores aquando do acesso ao site. Uma vez comprometidos os sistemas, os hackers utilizavam *backdoors* para estabelecer a comunicação entre as máquinas infectadas e os seus próprios centros de controlo. Esta comunicação era feita através de portos que estão associados a tráfego cifrado tal como a porta 443 (TCP). Uma vez com acesso a máquinas da rede interna das organizações, os hackers utilizavam as máquinas infectadas para aceder a outras máquinas e assim identificar outras vulnerabilidades e aceder a informações pertinentes. Posteriormente a informação era transmitida para os servidores utilizados pelos *hackers* para comunicar com as máquinas afectadas. A passagem de informação e reconhecimento da arquitectura das organizações aconteceu durante um longo período de tempo (Tankard, 2011).

Os casos acima descritos descrevem resumidamente a essência de alguns ataques para os quais qualquer organização deve estar alerta. Com a crescente adopção de dispositivos móveis para aceder à informação, a aumento da utilização de *botnets* assim como a sofisticação das chamadas "*Advanced Persistent Threats"* prevê-se que a ocorrência de ataques se torne ainda mais frequente.

Utilização de dispositivos móveis e suas aplicações

A utilização de dispositivos móveis *(Smartphones, PDA, etc.)* tem vindo a aumentar não só na vertente particular mas também na vertente profissional. A utilização destes dispositivos móveis a título profissional é de facto uma preocupação do ponto de vista de segurança dada a informação que estes dispositivos móveis contêm, quer sejam emails, dados de navegação, documentos, etc. (Simão et al., 2011). O facto de os dispositivos móveis serem cada vez mais utilizados faz com que cada vez mais aplicações sejam desenvolvidas para estes, tornando-os assim num alvo de ataque bastante atraente (Davi et al., 2011).

Por exemplo nos *SmartPhones* , uma grande parte destes tem como base o sistema operativo móvel *Android* que consiste num *kernel* baseado no *kernel* do Linux. Segundo um relatório elaborado pela McAfee, o *Android* tornou-se um dos alvos mais procurados pelos criadores de *malware* móvel. Na figura 1 podemos ter alguma visibilidade de ameaças novas identificadas no segundo trimestre de 2011 com base em malware móvel:

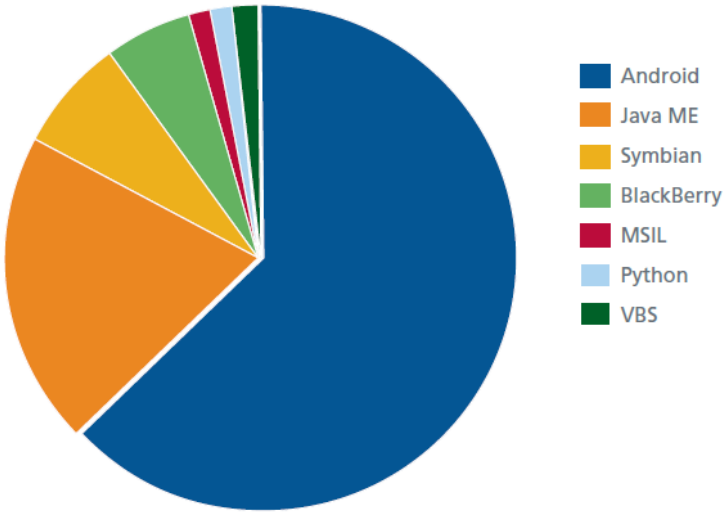


Figura 1: Novas ameaças móveis no 2º trimestre 2011 (adoptada de (McAfeeLabs, 2011)**)**

Existência das *Botnets*

Resumidamente, uma *botnet* é uma rede formada por máquinas infectadas por *bots* sendo que um *bot* é definido como sendo *software* malicioso que permite que terceiros tenham um controlo das máquinas infectadas por este *software*.

A utilização das *botnets* tem vindo a ser uma prática comum no domínio do *cybercrime* dada a potencialidade que estas possuem para o roubo e manipulação de informação das organizações. Para além disso, as *botnets* podem ser utilizadas para ataques de *Distributed Denial of Service* (DDoS) que podem causar perdas enormes principalmente em organizações cuja sua actividade

depende de serviços Web. A detecção de máquinas infectadas por *bots* e a monitorização das suas actividades nem sempre é uma tarefa fácil para os profissionais de segurança das organizações. Actualmente existem no mundo centenas de *botnets* conhecidas que são consideradas como uma constante ameaça que requer uma protecção contínua (Plohmann, Padilla, & Leder, 2011).

Advanced Persistent Threats

As *advanced persistent threats* são consideradas as principais classes de ameaças do momento. A sua designação está relacionada com as características destas ameaças, isto é, são utilizadas técnicas avançadas por parte dos atacantes para infectar máquinas identificadas como alvo e o tempo de recolha de informação por parte dos *malwares* é bastante dilatado.
Ultimamente, alguns dos grandes ataques com sucesso como foi o caso da operação aurora, do ataque à RSA Security e do ataque à Sony têm utilizado este tipo de filosofia e uma vez que os resultados foram positivos para os atacantes a tendência á que novas descobertas se venham a fazer.

Segundo vários autores, as tecnologias de segurança, tal como *firewalls*, detectores de intrusões, antivírus, etc., utilizadas tradicionalmente nas organizações começam a deixar de ser suficientes por si só para identificar e prevenir da melhor forma os constantes e sofisticados ataques (Tankard, 2011). A adopção dos sistemas de *Security Information and Event Management* (SIEM) começa a ser cada vez mais uma aposta das organizações, sendo que estes sistemas não vêem de forma algum substituir as tecnologias existentes, mas sim aproveitar informação recolhida dessas diversas tecnologias para identificar padrões que de outra forma passam despercebidos (Binde et al., 2011).

Problema Estudado e Objectivos

Os sistemas SIEM (*Security Information and Event Management*) têm vindo a assumir uma importância crescente na investigação que é feita em torno da detecção de intrusões e comportamentos anómalos nas redes e nas infra-estruturas tecnológicas, realizada quer em Universidades quer pelas empresas que fornecem este tipo de sistemas. Os desenvolvimentos que estes sistemas têm vindo a integrar na última década, resultam de uma necessidade do mercado em proteger os seus sistemas, cada vez mais complexos e mais expostos a ataques

externos e internos devido às mais variadas possibilidades de acesso através da Internet bem como aos inúmeros acessos de utilizadores autorizados.

Assim, são cada vez mais as empresas que recorrem a estes sistemas para garantirem a segurança da informação que gerem sendo este um aspecto particularmente importante para assegurar a confiança dos seus clientes. Devido ao custo elevado dos SIEM bem como complexidade da sua implementação, torna-se necessário torná-la eficaz, tanto do ponto de vista do tempo de execução das actividades de implementação como da eficiência dos mecanismos de detecção de ataques que estes sistemas permitem.

Os profissionais desta área necessitam de informação detalhada sobre a melhor forma de executar a implementação bem como orientações para resolver problemas específicos de determinadas configurações tecnológicas. No entanto, em resultado da experiência do aluno que trabalhou durante um ano como consultor numa empresa que faz implementações de tecnologias SIEM, bem como em resultado da revisão da literatura, concluiu-se que não existe muita informação que ajude o profissional desta área a resolver os problemas concretos com que se defronta no seu dia-a-dia.

Em particular, parece particularmente relevante o apoio no que concerne às decisões de implementação relativas à visibilidade da utilização de infra-estrutura IT e identificação de possíveis actividades anómalas, numa perspectiva de segurança de informação, através de utilização de sistemas SIEM.

Tendo em conta estas limitações da informação actualmente disponível ao implementador, com este trabalho propôs-se:

- Fazer uma pesquisa bibliográfica sobre métodos existentes de implementação de SIEMs;
- Analisar casos reais de implementações actualmente terminadas, em Portugal;
- Identificar métricas para avaliar implementações;
- Propor um novo método de implementação de SIEM com base em resultados de experiências práticas;

Estes objectivos foram sendo alcançados durante os 10 meses em que o projecto foi realizado. O primeiro objectivo foi alcançado pela revisão de literatura presente neste documento. Esta revisão de literatura permitiu analisar artigos científicos que descrevem avanços na área de sistemas de monitorização, como por exemplo a correlação de eventos de diversas fontes e com formatos diferentes. O segundo objectivo foi sendo alcançado ao longo do estágio profissional

que o aluno fez na empresa Unisys Portugal, onde analisou e acompanhou algumas implementações de SIEM.

No que diz respeito às métricas, foram identificados dois tipos de métricas sendo estas métricas consideradas técnicas ou tecnologias e métricas analíticas.

Uma vez que a literatura que descreve todo o processo de implementação é escassa, é objectivo deste trabalho contribuir nessa área pelo desenvolvimento e apresentação de um método de implementação. A descrição deste método é feita no capítulo IV e integra a descrição das várias fases do mesmo, onde se inclui a referência às principais preocupações que implementador deve ter em cada fase.

Metodologia de Investigação

Para que os objectivos desta dissertação fossem devidamente alcançados recorreu-se a duas metodologias de investigação passando estas por um *survey* bibliográfico e pela utilização do método *Design Research*. Seguidamente é apresentada uma breve descrição da utilização de cada uma deles:

Survey Bibliográfico

Para este estudo foram analisados artigos e livros sobre implementação de sistemas de detecção e monitorização de intrusões, contendo informação detalhada sobre os vários tipos de sistemas bem como sobre as melhores práticas que têm vindo a orientar a sua implementação.

A revisão de literatura foi efectuada em três fases: recolha de documentos, classificação de documentos por tema abordado, e agrupamento dos artigos que abordam as práticas de implementação de SIEM.

A tabela 1 apresenta uma visão global dos documentos recolhidos e analisados:

Tipo de Documento	Quantidade
Artigo em Revista Científica	7
Artigo em Conferência	5
White Paper	5
Relatório de Entidade relevante	5
Livro	4
TOTAL	26

Tabela 1: Visão Global dos Documentos Recolhidos e Analisados (Tipo)

A tabela 2 apresenta a informação relativa ao número de documentos em intervalos de tempo. Considerou-se que no período entre 2009 e 2011 estão os documentos que referem as tendências mais recentes na área estudada:

Ano	Nº Artigos
Anterior a 2000	3
2000 – 2008	13
2009 – 2011	10
TOTAL	26

Tabela 2: Documentos por Intervalo de Tempo

Depois de recolhidos os documentos, estes foram separados por tipo de sistema abordado, obtendo-se a classificação descrita na tabela 3:

Tipo de Sistema	Nº Artigos	%
IDS	16	62 %
SIEM	10	38 %
Práticas de implementação	14	54 %

Tabela 3: Documentos por tipo de sistema abordado

A análise dos artigos recolhidos permitiu determinar a evolução sistemas de segurança e monitorização de intrusões e a relevância que estes sistemas têm vindo a assumir nas práticas de segurança adoptadas pelas empresas. Simultaneamente foi possível também validar a relevância de desenvolver um método para orientar o processo de implementação de SIEM, proposta deste trabalho de investigação.

O método de *Design Science Research*

Ao longo do presente capítulo desde trabalho foi sendo identificada a necessidade de utilização de tecnologias mais sofisticadas, que utilizem como variáveis de entrada o resultado de outras tecnologias de forma a encontrar mecanismos que permitam identificar determinado tipo de ataques, onde os métodos e as tecnologias habituais falham. De forma breve, neste trabalho pretende-se apresentar um modelo de implementação de SIEMs, resultante da experiencia do aluno em projectos de implementação deste tipo de sistemas, em Portugal.

Para criar o conhecimento necessário escolheu-se como método de investigação o *Design Science Research*. A figura 2 apresenta uma visão global do método.

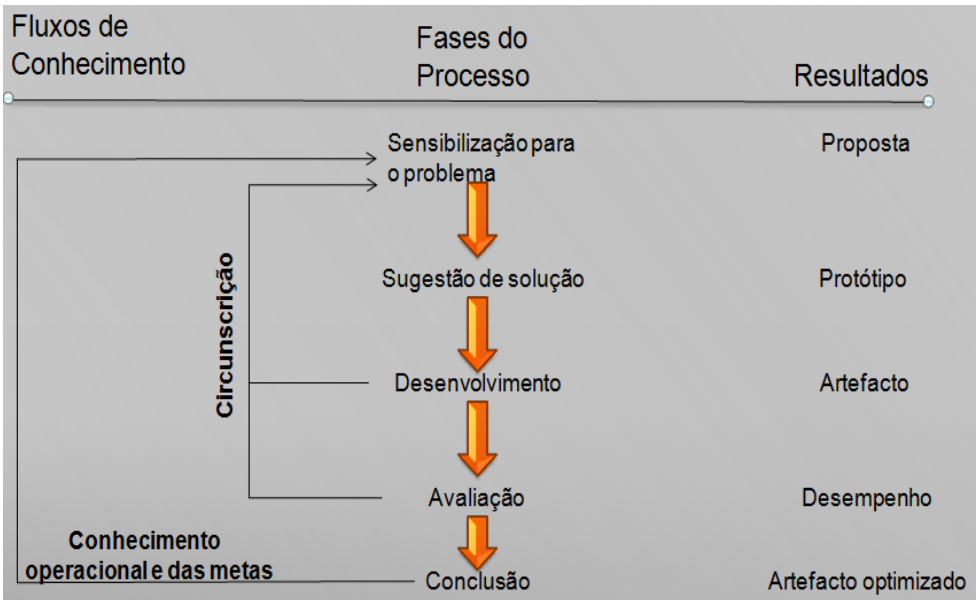


Figura 2: Método Design Science Research (adaptado de (Järvinen, 2007))

A figura apresenta como primeira fase do processo é "Sensibilização para o problema". Esta sensibilização decorre da constatação da existência de uma lacuna de conhecimento, a qual o investigador se sente motivado a preencher e que é possível de resolver dentro das restrições de tempo e outros recursos a que o projecto de dissertação está sujeito. O resultado desta fase é a descrição do tópico a investigar e do plano de trabalho a realizar para criar o conhecimento necessário. Considera-se que neste trabalho de dissertação, esta fase decorreu no período de preparação para a realização da dissertação e teve como resultado o documento denominado "projecto de dissertação".

A fase seguinte corresponde à "Sugestão da Solução". Esta fase corresponde a encontrar na literatura o essencial de uma solução para a resolver a lacuna de conhecimento e que serve de suporte teórico para o artefacto conceptual, metodológico ou tecnológico desenvolvido no projecto. No presente trabalho de dissertação, o Capítulo III contêm a descrição da literatura relevante na área de SIEM de onde se identificaram praticas de implementação. A partir destas práticas foram obtidas as principais actividades do processo de implementação a desenvolver na fase seguinte do método de investigação (figura 3).

Figura 3: Método de Implementação de SIEMs

Na fase de "Desenvolvimento", o artefacto é desenvolvido e implementado de acordo com as orientações emanadas da fase anterior. É nesta fase que a maior parte do trabalho de investigação é realizada, obtendo-se como resultado o artefacto completamente definido e pronto a testar. O Capítulo IV apresenta o método de implementação de um SIEM completamente definido. O Capítulo V apresenta a validação do método:

- Implementação de um SIEM numa empresa portuguesa aplicando o modelo definido;
- Verificação do resultado das métricas obtidas no final da implementação;
- *Survey* aos utilizadores dos SIEM.

De realçar que esta validação do método não foi feita neste trabalho por a fase "Avaliação" não poder ser enquadrada no período de tempo disponível para realização de uma dissertação de mestrado – um semestre. Encontrando-se o aluno integrado num projecto a realizar numa grande empresa Portuguesa e que terá início em Novembro de 2011, o aluno terá então oportunidade de fazer uma validação no contexto de uma empresa real e com uma elevada complexidade em termos dos equipamentos e comportamentos de utilização que tem necessidade de monitorizar.

Apesar da definição essencialmente sequencial das fases Desenvolvimento e Avaliação, os resultados obtidos podem conduzir a redefinições no entendimento prévio sobre a lacuna de conhecimento a colmatar. Até ao momento, tal redefinição ainda não foi sentida como necessária.

A fase de Conclusão termina o ciclo de concepção do artefacto, determinando a qualidade e impacto do artefacto produzido e terminando o projecto. Com a conclusão do projecto, a lacuna de conhecimento deverá ter sido colmatada, eventualmente criando novas perspectivas sobre o

tópico estudado que levem à identificação de novas lacunas ou à necessidade de desenvolvimento de outros tipos de artefactos.

Organização do Documento

Nesta secção é feita uma descrição de como se encontra organizado este documento assim como uma pequena síntese dos capítulos que o constituem.

Neste primeiro capítulo, para além da descrição da organização do documento, encontram-se descritos alguns casos problemáticos de ataques informáticos como forma de enquadramento e que levaram à motivação para o desenvolvimento desta tese. Para além disso encontra-se uma secção com o principal problema estudado e os objectivos esperados ao que se seguiu uma descrição sintética da metodologia de investigação utilizada ao longo do semestre.

No capítulo II procede-se com uma revisão de literatura relacionada com os sistemas de detecção de intrusões, sendo que estes são considerados como sendo a base dos SIEM. É facultada uma definição do conceito segundo vários autores, seguindo-se uma descrição das principais categorias existentes assim como alguns exemplos de colocação destes sistemas nas redes das organizações. No decorrer do capítulo para além de referência aos principais métodos de detecção de intrusões são apresentadas algumas soluções existentes neste tipo de sistema e enumeradas algumas das principais limitações. Por fim são tidas em conta algumas considerações sobre o assunto que explicam a necessidade da evolução dos sistemas de detecção de intrusões.

No capítulo III é também relacionado com a revisão de literatura, desta feita relacionada com os sistemas SIEM. O capítulo começa com uma pequena explicação da origem do SIEM assim como a definição do conceito, da arquitectura genérica e das principais funções destes sistemas Posteriormente é descrita uma visão geral da sua evolução ao longo do tempo, são apresentadas possíveis métricas e identificadas algumas soluções existentes no mercado.

O capítulo IV é inteiramente dedicado à descrição do método de implementação de SIEMs, desenvolvido ao longo do semestre com base da experiencia do aluno em projectos de implementação de SIEMs em Portugal, ao cargo da empresa Unisys. Após uma breve descrição da figura que resume o método desenvolvido é apresentada uma breve descrição de cada uma das cinco fases, assim como a definição e exemplos práticos de cada tarefa identificada para cada fase.

Já no capítulo V são apresentadas duas medidas de validação do método proposto, sendo que uma delas passa pela aplicação num projecto real e a segunda passa pela apresentação de métricas retiradas de um outro projecto.

Por fim, no capítulo VI são apresentadas as principais conclusões do trabalho que foi feito, tal como as principais contribuições que este acrescenta na área da segurança dos sistemas de informação. Para além disso são enumeradas algumas limitações inerentes ao trabalho que foi desenvolvido mas também algumas reflexões sobre oportunidades identificadas para investigações futuras.

Capítulo II - Sistemas de Detecção de Intrusões

Conceito

O conceito de *Intrusion Detection System* (IDS), também conhecido em Portugal como "Sistema de Detecção de Intrusões", é um conceito de que se fala há mais de 20 anos. Já em Fevereiro de 1987, Denning abordou o conceito de *Intrusion Detection* num artigo denominado de "*An Intrusion-Detection Model*". Neste artigo o autor identifica alguns factores que justificam a existência de um sistema de detecção de intrusões em "tempo real" focando as vulnerabilidades que as soluções da altura apresentavam e alguns problemas que se pretendiam resolver, relacionados com ataques e utilização abusiva de privilégios por parte de utilizadores internos (Denning, 1987).

Numa publicação do *National Institute of Standards and Technology* (NIST) em 2001, Barber e Mell definem *Intrusion Detection Systems* como sendo sistemas de *hardware* ou *software* que automatizam o processo de monitorização de eventos que ocorrem em sistemas e redes de computadores, ajudando a analisar os sinais de problemas de segurança (Barber & Mell, 2001). Ao longo dos anos o conceito de IDSs tem-se mantido, apenas com pequenas alterações que não afectam o cerne do significado do conceito. Prova disso é a definição apresentada em 2009 por Lu Hong que define *Intrusion Detection Systems* da mesma forma que Barber e Mell, clarificando apenas quais os problemas de segurança que este tipo de *software* detecta, com vista a evita-los e com foco no uso indevido de sistemas e redes de computadores por parte de utilizadores internos nas organizações (Hong, 2009).

Categorias

O facto deste conceito se manter estável ao longo dos anos revela a importância que estes sistemas mantêm nos dias de hoje. No entanto tem havido evoluções ao nível da tecnologia que o implementa, nomeadamente com o aparecimento de vários tipos de IDSs tal como o *Host-Based* IDS (HIDS), o *Network* IDS (NIDS) e o *Wireless* IDS. No entanto as principais referências continuam a ser os *Host-Based* e os *Network* IDS. A principal distinção entre estas várias formas de implementação está relacionada com o tipo de análise que é feita; mais concretamente, segundo a definição apresentada num relatório publicado pela *Information Assurance Technology Analysis Center* (IATAC), os *Host-Based* IDS são os sistemas responsáveis por analisar configurações específicas num sistema, tal como acesso a um determinado *software* ou

políticas de seguranças locais; a análise é feita ao nível das próprias máquinas. Os *Network* IDS são definidos como sendo sistemas de análise de tráfego nas redes ao nível de todas as camadas da *Open Systems Interconnection* (OSI), tomando decisões acerca da finalidade do tráfego e analisando actividades suspeitas (Tyler & Wu, 2009); a análise é feita com base nas redes e no tráfego que estas suportam.

Tal como o próprio nome indica, um IDS tem uma funcionalidade principal que é a detecção de intrusos ou detecção de comportamentos indevidos num sistema. Para que isto possa acontecer, os IDSs estão preparados com um conjunto de funcionalidades que permite aumentar a eficácia destes sistemas na execução das suas tarefas. Num artigo de Dennis Mathew, publicado em 2002 pelo SANS Institute, podem-se identificar seis funcionalidades claras que os IDSs possuem para atingir o principal objectivo, sendo que estas funcionalidades passam pela monitorização de utilizadores e sistemas, análise de configurações dos sistemas assim como as vulnerabilidades a estes associados, avaliação dos sistemas de arquivo assim como a sua integridade, capacidade de detecção de padrões de ataques já conhecidos e controlar as violações da política de segurança por parte dos utilizadores (Mathew, 2002).

De forma a fazer uma pequena análise às funções no que diz respeito à monitorização, podemos verificar na tabela 5 que os diferentes tipos de IDSs acima referidos oferecem vários tipos de funções. No entanto, apenas a identificação de violações da política de segurança é feita pelos dois tipos de IDSs.

	IDS Category	
	Host-IDS	Network-IDS
Identificar acessos não autorizados a ficheiros e outros recursos do sistema.	X	
Identificar violações da Politica de Segurança.	X	X
Identificar "*Trojan Horses*" e *software* malicioso.	X	
Identificar ataques a serviços de rede.		X
Identificar ataques de negação de serviço (DoS).		X
Identificar falhas ou má configuração nas *firewalls*.		X
Identificar ataques a redes cifradas ou comutadas	X	

Tabela 4: Diferentes tipos de funções de monitorização (adoptada de (Andress, 2004))

Apesar de um sistema de detecção de intrusões ser um sistema importante do ponto de vista da segurança, a integração dum sistema deste tipo acarreta alguns desafios que têm de ser ultrapassados. Um destes desafios de implementação é conseguir fazer a gestão dos grandes volumes de tráfego que são gerados nos dias de hoje numa organização. Este facto acaba por remeter para um outro desafio relacionado com os falsos positivos e consequentemente com aumento de trabalho do ponto de vista dos administradores. Isto é, à medida que o tráfego aumenta, o número de falsos positivos gerados pelos IDSs tende a aumentar e estes têm de ser avaliados quanto à sua veracidade. Tal como explica Daniel Ragsdale *et al*. num artigo publicado no ano 1999, um IDS não é perfeito e gera falsos positivos e falsos negativos, fazendo em seguida questão de esclarecer que depois de cada (potencial) incidente reportado, o administrador do sistema deve indicar se este foi considerado um ataque ou se foi um falso alarme (Ragsdale et al., 1999).

Arquitectura de rede com detectores de Intrusões

A forma como os IDSs vão ser colocados na rede ou sub-rede da organização está relacionado com o tipo de detecção e monitorização que se quer implementar.

A fronteira entre a rede interna e a rede externa é normalmente um dos pontos que é monitorizado de forma a identificar ataques direccionados à própria organização. No entanto, este tipo de monitorização faz com que o número de alertas seja bastante volumoso. Para que este tipo de monitorização seja possível, é necessário colocar um NIDS entre a *firewall* de perímetro e a Internet (IDS 1 na figura 4) ou colocar um sensor na própria *firewall*.

Outra prática bastante utilizada é a colocação de NIDS em cada sub-rede da organização de forma a garantir que estas são devidamente monitorizadas. Na figura 4 pode-se verificar a colocação do IDS 3 e do IDS 4 que estão a efectuar detecções sobre os *Hubs* à entrada de cada sub-rede.

Para além das duas praticas mencionadas anteriormente, é também comum a colocação de NIDS nos *switches* que fazem a divisão da rede em sub-redes. Este NIDS está representado na figura 4 pelo IDS 2.

Qualquer organização é detentora de *hosts* mais críticos ou *hosts* que requerem um cuidado redobrado. Nessas situações há possibilidade da colocação de agentes de HIDS (ver figura 4) de forma a identificar intrusões que os NIDSs não conseguem identificar, como o caso especifico de acesso a ficheiros por exemplo (Andress, 2004).

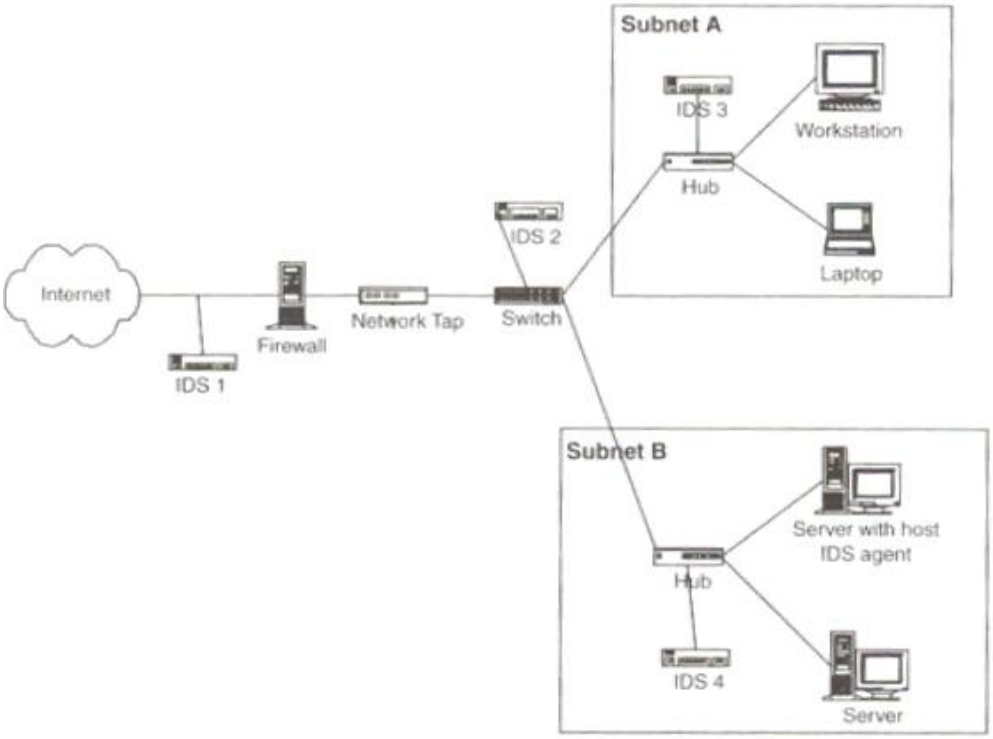


Figura 4: Posicionamento dos IDSs na Rede (adaptada de (Andress, 2004))

Métodos de detecção

A detecção de intrusões pode ser feita por dois tipos diferentes: anomalias e assinaturas. No caso dos IDSs baseado em anomalias, os sensores são configurados para detectarem actividades suspeitas com base nos padrões normais. Por exemplo, se o consumo da largura de banda aumentar demasiado de acordo com o que é normal, se é identificado um aumento significativo da utilização de um determinado protocolo ou porta, se determinados pacotes de tráfego apresentam tamanho acima do normal, etc. A preparação de um IDS deste tipo requer algum tempo uma vez que este tem de identificar uma base do padrão normal dos comportamentos na rede, ou seja, o IDS tem que "aprender" os comportamentos aceitáveis para que o número de falsos positivos seja menor. Esta técnica de detecção tem associada uma mais-valia importante que é a capacidade de detecção de um ataque sem que haja um conhecimento antecipado do mesmo. Esta mais valia pode por vezes tornar-se um problema se os utilizadores dentro das redes pontualmente tiverem um comportamento que não é normal, mas que pode não ser malicioso, o que vai fazer com que o sensor do IDS identifique como sendo uma intrusão, não passando de um falso positivo (Barber & Mell, 2001).

No caso dos IDSs baseados em assinaturas, os sensores estão configurados para identificarem tráfego previamente conhecido como sendo tráfego malicioso. Este tipo de IDS carece de

constante actualização uma vez que apenas actua com base em acontecimentos já conhecidos, isto é, o administrador do IDS tem de constantemente actualizar a base de dados de assinaturas, quer para inserir novas quer para remover assinaturas que apenas estão a consumir recursos e, potencialmente, a conduzir a falsos alarmes. Os detectores de intrusões baseados em assinaturas geram um número reduzido de falsos positivos uma vez que apenas detectam intrusões conhecidas. No entanto, apesar de ter a base de dados de assinaturas actualizada, o IDS perde alguma eficiência uma vez que ataques derivados de ataques já reconhecidos, desde que produzam assinaturas diferentes, não serão identificados como intrusão (Hofmeyr et al., 1998).

Como seria de esperar e atendendo à sua complementaridade, é possível (e desejável) ter numa mesma solução, ambas as técnicas de detecção, numa arquitectura híbrida. Na Figura 5 podemos verificar um diagrama que representa a arquitectura genérica de um IDS híbrido.

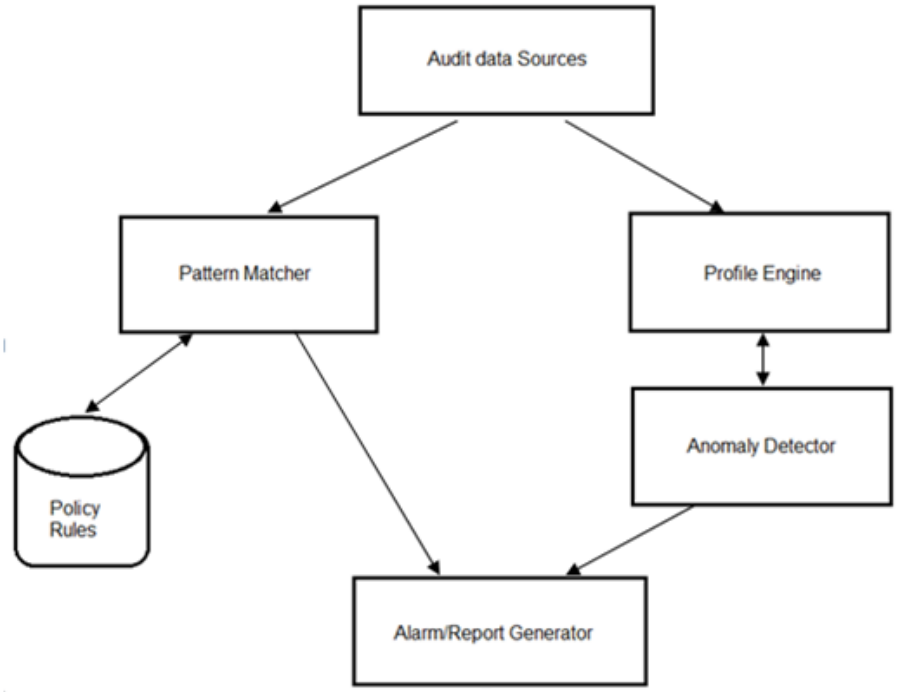


Figura 5: Arquitectura genérica de um IDS híbrido (adoptada de (Bace, 2000))

Quanto ao mecanismo (algoritmos) de detecção os IDS podem recorrer a diferentes técnicas. Algumas destas técnicas estão mais associadas a um método de detecção e outras mais associadas a outro. Os IDSs baseados em anomalias utilizam técnicas como *rule modelling, descriptive statistics, artificial neural network, simple stat, simple rule-based, threshold* e *state series modelling.* Já os IDS baseados em assinaturas utilizam técnicas como *state-transition* e *petri-net* (Axelsson, 2000).

Implementações mais conhecidas

O mercado dos sistemas de detecção de intrusões é bastante grande dada a importância dos mesmos e devido ao seu tempo de existência e maturidade. Um dos IDS mais conhecido e utilizado no mercado é o *Snort*, quer pelas suas capacidades e características quer por ser um *software open source* (Andress, 2004). Para além deste existe uma panóplia de outros sistemas tais como: *HP-UX HIDS, IBM Real Secure Server Sensor, McAfee Host Intrusion Prevention, NetIQ Security Manager iSeries, OSSEC HIDS, Tripware, Cisco Catalyst, Cisco Guard, Enterasys Dragon Network Defense, Imperva SecureSphere, Sourcefire, TippingPoint* (Tyler & Wu, 2009).

Principais Limitações

Uma das limitações dos IDSs é a dificuldade que estes apresentam em detectar intrusões em redes de grande dimensão com *firewalls* e *bastante segmentada*. Os dispositivos de rede podem ainda usar diferentes formatos de dados e protocolos de comunicação, os quais devem ser reconhecidos pelo IDS para que este possa detectar intrusões de forma eficaz. Isto apesar de alguns esforços na implementação de IDS de forma distribuída, mas cujos resultados evidenciam limitações, para além de uma elevada dificuldade na implementação(Huang et al., 1999).

Quanto maior for a rede, mais são as vulnerabilidades que podem ser exploradas pelos *hackers* já que se torna muito difícil ao IDS monitorizar todo o tráfego. Mesmo se forem colocados vários componentes de NIDS ao longo da rede, haverá sempre necessidade de assegurar que se encontram em pontos estratégicos para que os ataques não iludam os sensores utilizando percursos alternativos.

O sistema operativo pode também dificultar a detecção de intrusões devido às vulnerabilidades que integra e que podem ser exploradas pelos *hackers*. O IDS pode simplesmente não detectar que o sistema operativo foi alterado de forma ilegítima e, portanto, não detectar actividade anómala.

Os NIDSs podem não detectar tráfego na rede quando estão a ser usados VPN e *scured encrypted tunnels* para mascarar intrusões. Nestes casos, a menos que o NIDS saiba decifrar e voltar a cifrar os dados, os ataques podem passar despercebidos. Também à medida que o tráfego aumenta, integrando muitos canais de comunicação, fica mais difícil analisar os dados com rapidez e acuidade suficiente. Em redes muito grandes, as intrusões podem iludir os sensores do NIDS. No caso das redes sem fios, torna-se relativamente simples interceptar os

data packets, no entanto estas redes utilizam protocolos muito próprios. Os IDSs têm que compreender também esses protocolos e isso pode ser um desafio. Por outro lado, se um sistema ligado a uma rede *wireless* já estiver comprometido devido a um ataque anterior, ao juntar-se à rede pode ser difícil ao IDS detectar intrusões de um sistema identificado como fiável (Tyler & Wu, 2009).

Considerações Finais

Os IDSs podem não ser capazes de detectar ataques provenientes de múltiplas fontes pelo que surge a necessidade da utilização de sistemas integrados que possam relacionar actividade intrusiva em vários pontos da rede. Sistemas de detecção de intrusões são uma parte importante da segurança da infra-estrutura. No entanto, por si só podem não ser suficientes para uma detecção eficaz. A correlação dos eventos gerados pelos vários tipos de IDS assim como a correlação destes com eventos de outros equipamentos como *firewall* ou *routers* tornou-se numa mais-valia para a área de segurança (Tyler & Wu, 2009). Por isso, o administrador da rede deve utilizar outras ferramentas em complementaridade (Barber & Mell, 2001).

Estas limitações levam a que se tenha sempre em atenção alguns aspectos quando se procura determinar qual a melhor solução. Tal como indica Amanda Andress, "um dos aspectos que não pode ser esquecido é o grau de precisão com que os IDS identificam os ataques ou as actividades suspeitas" (Andress, 2004). No entanto, este tipo de avaliação é extremamente difícil de precisar uma vez que a variedade de ataques é bastante grande e as formas como estes ataques são despoletados são ainda mais numerosas.

De forma a colmatar as principais fragilidades e limitações dos IDS e de outros equipamentos de segurança que por si só não eram suficientes, no mercado apareceram implementações designadas de SIEM (*Security Information and Event Management*) cujo objectivo é centralizar os eventos gerados pelos equipamentos das infra-estruturas, correlaciona-los entre si de forma a identificar ataques ou actividades maliciosas e gerar alertas quando identificado algum possível ataque. Estes esforços têm sido acompanhados por algum trabalho científico ao nível de técnicas de correlação de eventos de segurança e *logs*, mas cujos resultados ainda estão aquém de quantitativamente indicarem uma solução prática (Myers et al., 2011).

Capítulo III - Security Information and Event Management

Origem

O conceito de *Security Information and Event Management* (SIEM) tal como o nome indica, resulta da fusão dos conceitos de *Security Information Management* (SIM) com o conceito de *Security Event Management* (SEM). Quer os SIM quer os SEM, ambos focam os mesmos aspectos relacionados com a recolha e análise de eventos de segurança. No entanto, enquanto os SEM têm uma grande capacidade de agregação de grandes volumes de dados, estando mais voltados para a análise e detecção de incidentes de segurança em tempo real fazendo uso de motores de correlação, os SIM são utilizados com objectivo de manter um histórico de eventos da infra-estrutura durante longos períodos de tempo (Gabriel et al., 2009).

Na figura 6 podemos verificar uma arquitectura conceptual dos SIEM onde estão representadas características próprias destes sistemas assim como características herdadas que resultaram da fusão entre os SEM e os SIM.

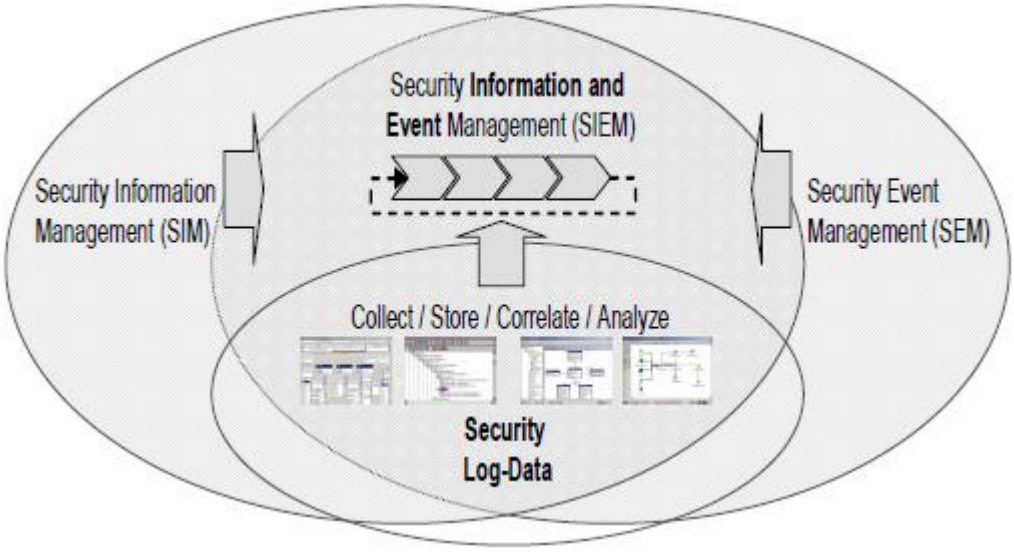


Figura 6: Arquitectura conceptual dos SIEM (adoptada de (Gabriel et al., 2009)**)**

Definição

Security Information and Event Management (SIEM) é um tipo de solução que permite a gestão centralizada de eventos de diversas fontes, assim como a correlação desses mesmos eventos. Normalmente são compostos por dois módulos principais: um módulo para armazenamento

proveniente dos sistemas SIM e outro módulo para efectuar análise e investigações sobre os eventos proveniente dos sistemas SEM. A recolha dos eventos para os sistemas de SIEM pode ser feita por dois métodos distintos, nomeadamente *Agentless* (sem recurso a agentes) ou *Agent-Based* (utilizando agentes). Pelo primeiro método, *Agentless*, os eventos são enviados periodicamente pelos equipamentos autenticados para servidores específicos, ou o servidor autenticado vai buscar os eventos directamente a esses equipamentos. Com os eventos no servidor, é então possível fazer a sua análise, a qual pressupõe tarefas de filtragem, agregação e normalização para que possam ser detectadas actividades relevantes para o controlo de segurança que se pretende efectuar (Kent & Souppaya, 2006).

No caso do método de recolha de eventos por *Agent-Based*, esta é efectuada por meio de conectores, também designados por colectores ou agentes. Este *software* é instalado no equipamento sendo as tarefas de filtragem, agregação e normalização feitas localmente. Depois de realizadas essas tarefas, os eventos são enviados para um sistema SIEM, em tempo real, que faz o seu armazenamento e análise. Os conectores usados correspondem a determinados tipos de eventos. Se a recolha e análise feita por servidor englobar vários tipos de eventos (sistema operativo, aplicações, etc.), então são necessários vários conectores instalados nos respectivos equipamentos identificados para essa análise. Convém ainda realçar que os SIEM trazem conectores genéricos que tratam mais do que um tipo de evento. No entanto, enquanto alguns fabricantes apostam mais em fornecer os conectores pré-configurados para as tecnologias mais utilizadas no mercado, outros optam por fornecer aos administradores as condições para estes criarem os conectores necessários para o seu ambiente (Kent & Souppaya, 2006).

Cada um destes métodos (*Agentless* e *Agent-Based*) apresenta vantagens e desvantagens específicas. Os SIEM baseados em *Agentless* são mais simples de instalar mas são menos eficientes na filtragem de eventos tornando a análise um processo mais complexo e eventualmente mais falível. Os SIEM considerados *Agent-Based* são mais complexos na instalação mas permitem um tratamento e análise dos eventos mais eficiente (Kent & Souppaya, 2006).

Arquitectura genérica dos SIEM

A figura 7 reflecte a descrição que foi apresentada anteriormente acerca do funcionamento dos SIEM.

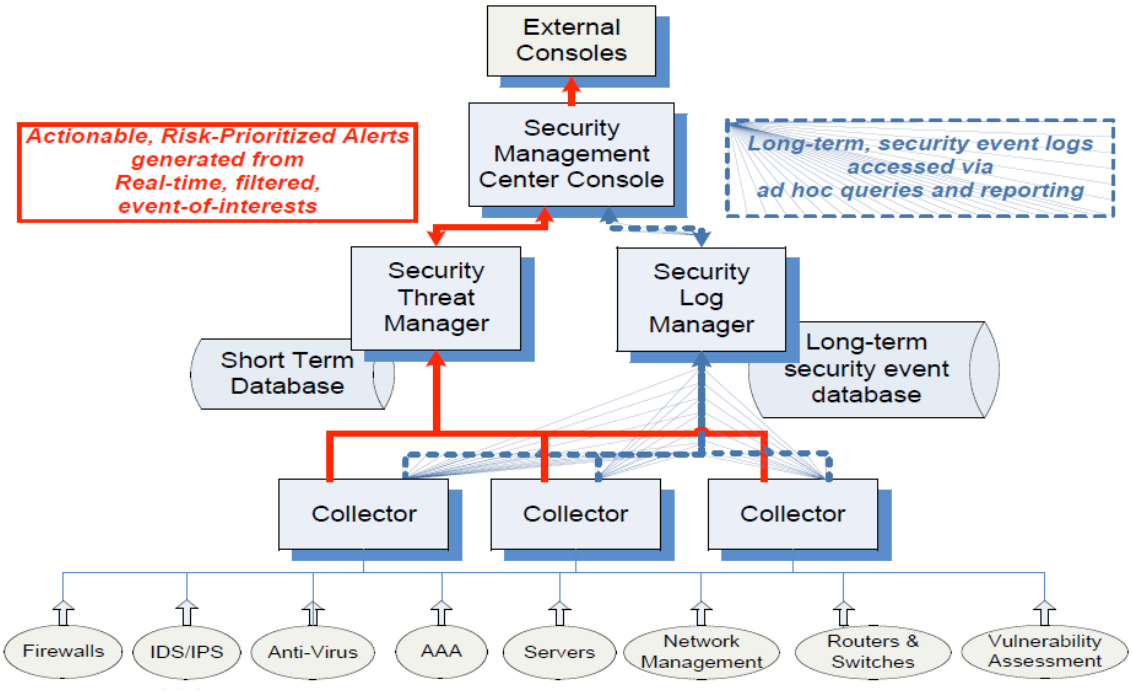


Figura 7: Arquitectura genérica dos SIEM (adoptada de (Swift, 2006))

Resumidamente, verificamos na camada mais inferior da figura uma representação dos equipamentos de segurança das organizações, com capacidade para gerar eventos. De alguma forma esses eventos chegam aos colectores ou conectores através de algum protocolo. Posteriormente estes conectores processam os eventos (normalização, filtragem, agregação, etc) e com base em configurações previamente definidas encaminham os eventos para a componente inteligente onde é feita correlação (através do fluxo representado a vermelho na figura), para a componente de arquivo também designada por *log management* (através do fluxo representado a azul na figura), ou para ambos em simultâneo mediante o evento em causa. Para além disso, posteriormente estes eventos podem ser analisados de forma centralizada através de uma consola de gestão, que tanto pode estar local como remota, podendo ainda estar associadas a sistemas de gestão de incidentes.

Principais Funções dos SIEMs

Figura 8: Principais funções dos SIEM

Tal como se pode verificar na figura 8, os SIEM têm quatro tarefas principais. O sucesso de cada uma das fases vai depender de cada uma das fases imediatamente anterior, isto é, analisando a figura podemos concluir o processo de consolidação dos eventos apenas obterá os resultados esperados se o processo de recolha de eventos correr na perfeição, assim como a correlação de eventos apenas será bem-feita se estes já obedecerem aos requisitos que os processos anteriores exigiam e assim sucessivamente.

Recolher Eventos

Esta primeira fase de recolha de eventos é a que está na base de toda a ferramenta SIEM. Sem esta camada de recolha na arquitectura desta tecnologia, seria impossível extrair informações sobre eventos decorridos e, consequentemente, não existiria qualquer gestão de segurança da rede e toda a funcionalidade de uma ferramenta SIEM deixaria de ter sentido.

A recolha de eventos pelo método de Agentless utiliza o próprio software que faz parte do cerne dos SIEM. Tal como foi enunciado anteriormente, este método não é o mais indicado por motivos de degradação de desempenho e eficiência, portanto apenas deve ser utilizado para a recolha de eventos em quantidades baixas.

Já o processo de recolha de eventos por Agent-Based, baseia-se num agente que está associado a cada dispositivo da rede a monitorizar (seja um nó físico, uma *firewall*, um Sistema de Detecção/Prevenção de Intrusões, etc.). Esse agente (físico ou uma aplicação a ser executada em *background*) será configurado para activamente monitorizar um ficheiro recolhendo todos os novos eventos que derem entrada nesse ficheiro, em operações em bases de dados (por exemplo) e, a cada intervalo de tempo (configurável), enviará esses eventos para um servidor central, alimentando assim todo o sistema SIEM.

Estes agentes, para além da função de recolherem eventos, têm uma outra funcionalidade imprescindível. Porque cada aplicação/entidade da rede pode ter os eventos num formato próprio, seria árduo para o "cérebro" da plataforma SIEM conseguir deduzir toda a informação de todos os eventos nos mais diferentes formatos. Assim, a segunda função destes agentes é normalizar os eventos que está a escutar antes de os enviar para o servidor central. O objectivo deste processo de normalização dos dados é tratar os registos para que entrem na base de dados num formato comum e conhecido pelo sistema. Desta forma, será possível à plataforma retirar informação dos atributos de um evento muito mais facilmente.

Consolidar Eventos

Após o envio dos eventos para o servidor central, torna-se necessário ter esses eventos salvaguardados para a próxima fase de correlação dos mesmos, ou para futura auditoria. Tipicamente, numa solução SIEM, os eventos permanecerão consolidados numa Base de Dados central.

Guardar todos os eventos registados nos últimos tempos é vital não só para ter um registo actualizado do presente mas também para manter um histórico do passado, que pode ser de extrema importância na descoberta de padrões comportamentais que se notem no presente e que já foram sentidos no passado. O tempo que os eventos permanecem em base de dados está dependente do tipo de segurança que queremos implementar, mas tipicamente deve ser pela sua manutenção durante alguns meses.

Dada a importância inerente à informação dos eventos armazenados, é de uma importância extrema também eles serem alvo de alguma auditoria de segurança. Como tal, é altamente aconselhável aplicar sobre estes registos as típicas técnicas que salvaguardam os dados, como replicação da base de dados, cópias de segurança actualizadas, etc. Espera-se também que a informação contida na base de dados seja confidencial, pelo que habitualmente exigem-se métodos que garantam a confidencialidade e integridade da mesma.

Correlacionar Eventos

Agora que todos os eventos foram recolhidos e devidamente acondicionados numa Base de Dados central, seria possível, por exemplo, olhar para o passado e determinar se existiu algum tráfego que não devia ter atravessado uma *firewall*, violando assim as políticas de seguranças estabelecidas na organização. No entanto, para tirar todo o potencial de uma plataforma SIEM e de forma a extrair muito mais informação do que se passou na rede do que retirámos no exemplo anterior, é necessário correlacionar os eventos. Por exemplo, um evento registado numa *firewall* quando visualizado isoladamente pode não ter qualquer significado mas, provavelmente, se o correlacionarmos com outros eventos oriundos de um encaminhador, base de dados ou outro qualquer nó, poderá indicar um ataque a um sistema vulnerável.

O acesso a regras de correlação é uma tarefa relativamente simples uma vez que são disponibilizadas alguns excertos de código aberto que permitem identificar alguns ataques. No entanto, tal como indicam alguns autores do livro "SIEM Implementation" o facto de as

organizações sentirem necessidade de regras de correlação mais avançadas faz com que estas optem pela aquisição de SIEMs proprietários (Miller et al., 2011).

Na figura 9 temos um exemplo de uma regra de correlação simples em que é gerado um alerta sempre que forem identificados 3 ou mais eventos de tentativas de autenticação falhadas, prosseguidas de um evento de autenticação com sucesso, em que o endereço de origem e o endereço de destino seja o mesmo. Na figura não se encontra representado o período de tempo definido para o alerta ser gerado, no entanto este é um parâmetro a ter em conta aquando da definição das regras de correlação:

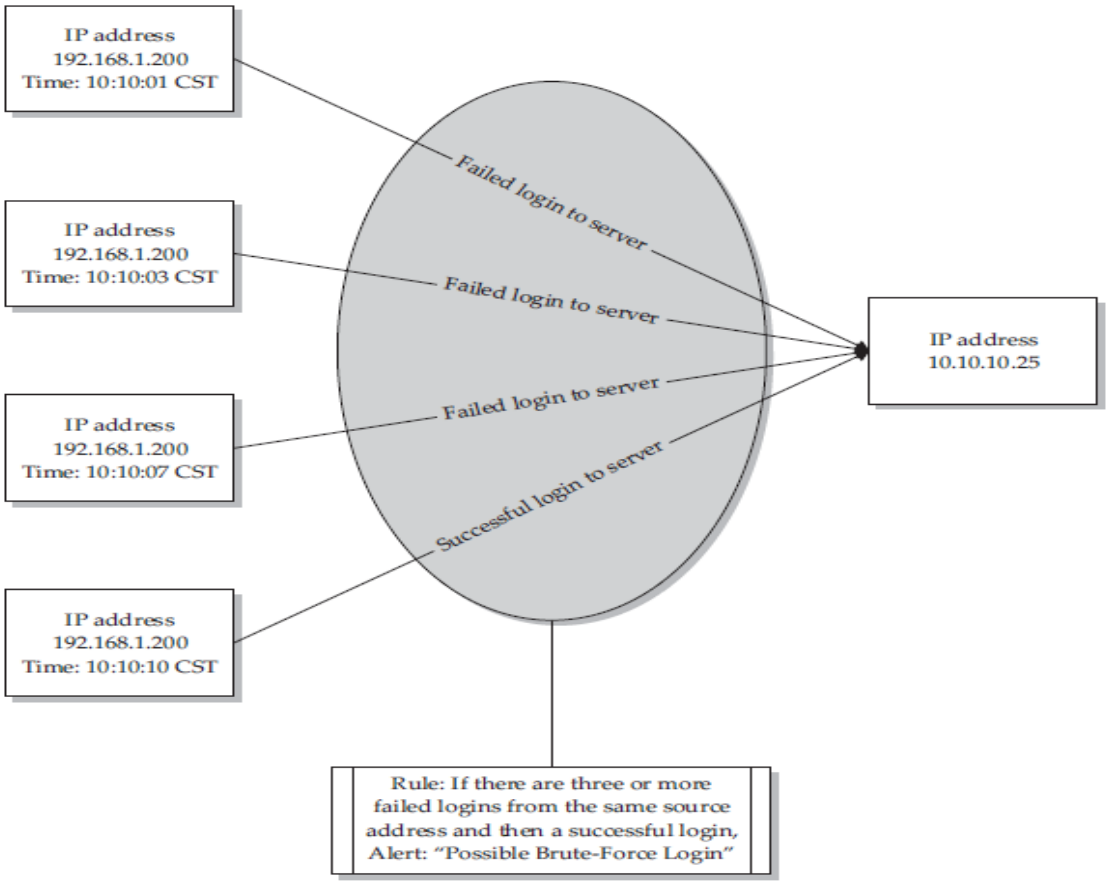


Figura 9: Exemplo de Regra de Correlação de *"Possible Brute-Force Login"* (Adoptada de (Miller et al., 2011)**)**

Na figura 10 temos um outro exemplo de correlação, desta feita com base em eventos de diferentes equipamentos, nomeadamente um IDS e uma *firewall*. Na figura o atacante encontra-se representado pelo quadrado azul enquanto a vítima pelo quadrado branco. Neste cenário a vítima tem associado um só endereço público, operando por via de NAT (*Network Adress Translation*). Na figura está representada a vítima mas com o seu endereço privado.

Na primeira fase do ataque, representado pelo número 1, o IDS gera um evento designado "Web-IIS ISAPI printer Access". Ao analisarem o evento, os analistas não conseguem identificar se o ataque foi bem sucedido ou não. Na fase 2, verificam-se eventos gerados pela *firewall* indicando que o atacante ultrapassou esta barreira, ainda assim os analistas não têm indicação de que o ataque foi bem sucedido ou não. Já na fase 3, a *firewall* identifica que a máquina interna atacada a partir do seu IP público fez algum tipo de comunicação com o endereço do atacante. Os motores de correlação têm a capacidade de correlacionar este tipo de eventos o que faz com que possam ser detectado ataques que não são identificados analisando os eventos individualmente.

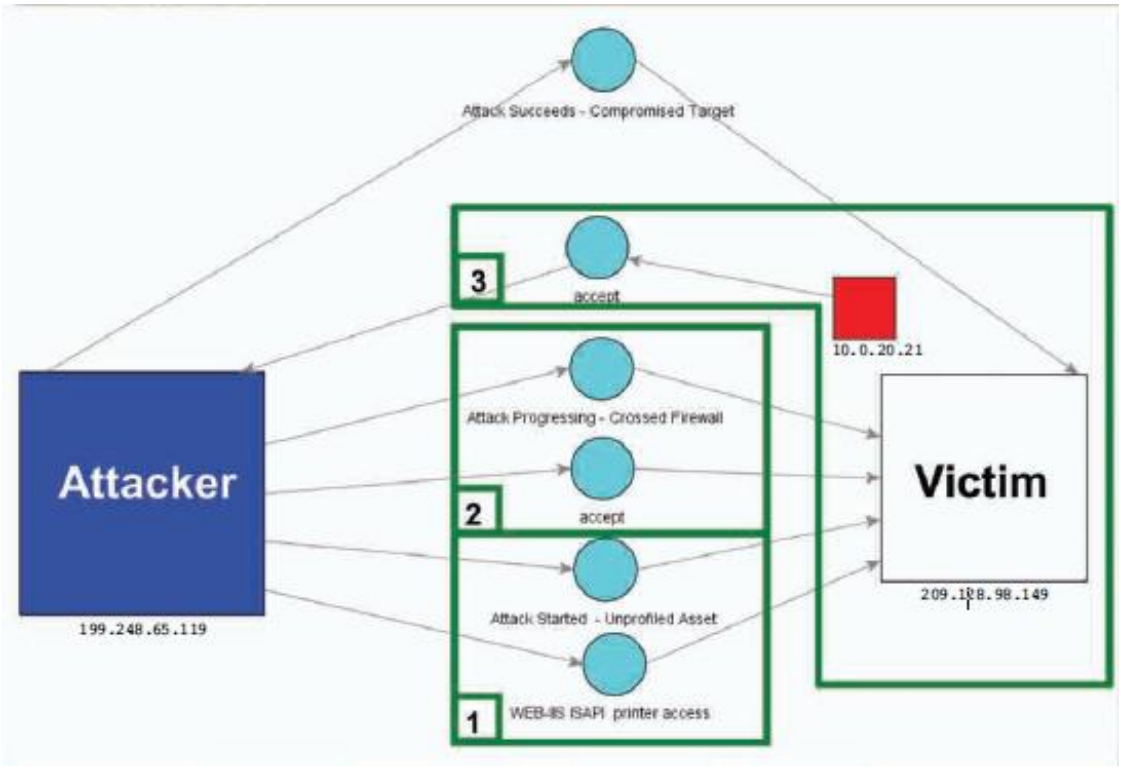


Figura 10: Exemplo de Regra de Correlação (Adaptada de (ArcSight, 2007)**)**

Esta fase de correlação é praticamente a fase final do ciclo de vida de um evento. A partir daqui, após estas relações estarem definidas, uma ferramenta SIEM permite tomar as mais variadas decisões.

Evolução

Apesar de o conceito de SIEM ser um tema actual no mercado, estes sistemas já existem á alguns anos, tendo passado por algumas fases de maturidade e evolução. Desde a geração 1.0, que focava a protecção de intrusões externas, passando pela geração 2.0, que acrescenta às

funcionalidades de um SIEM 1.0 funcionalidades de protecção da rede interna da organização, até aos SIEM 3.0 que, para além das características que herdaram do SIEM 2.0, ainda focam a protecção do negócio pela monitorização do comportamento dos utilizadores internos e externos à organização.

SIEM 1.0 – Protecção do Perímetro

A primeira geração de SIEM, designada SIEM 1.0, estava directamente associada ao conceito de protecção do perímetro. As *firewall*, os sistemas de detecção de intrusões (IDS) e os sistemas de prevenção de intrusões (IPS) eram os principais equipamentos a ter em conta de forma a manter protegida a rede interna das organizações, ou seja, eram a primeira linha de defesa contra a espionagem industrial, hackers ou qualquer tipo de malware – a isto se chama protecção do perímetro (Andress, 2004). A monitorização do perímetro e o cruzamento da informação gerada por esses equipamentos permitia que fosse identificado o bom funcionamento e configuração das tecnologias em causa, com base no tráfego que passava para a rede interna. Tendo em conta este tráfego era possível identificar possíveis máquinas comprometidas com necessidades de passagem para um estado de quarentena, ou seja, isolar essas máquinas para que se possam monitorizar as suas actividades e identificar possíveis infecções por *software* malicioso, por exemplo.

SIEM 2.0 – Protecção da Rede

A geração 2.0 dos sistemas SIEM provém de um acréscimo das funcionalidades dos sistemas SIEM utilizados na geração 1.0. Estes sistemas, para além das capacidades mencionadas anteriormente, resolvem as questões de conformidade que exigem o cumprimento de diversas normas de forma a garantir a protecção na rede interna. A informação relativa a antivírus e relativa às próprias máquinas começou a ser tida em conta nesta geração, permitindo assim identificar o estado de actualização dos antivírus, a forma como determinado vírus se espalha dentro da rede e quais as máquinas infectadas, para que se procedesse à sua reinstalação ou limpeza. Para além da monitorização do ciclo de vida de determinado vírus passou a ser possível também efectuar o *tracking* dos ataques efectuados á rede, com ou sem o sucesso pretendido.

SIEM 3.0 – Protecção do Negócio

A geração 3.0 dos sistemas de SIEM está relacionada com os sistemas de SIEM que têm a capacidade de recolher e correlacionar informação de todo o negócio de uma qualquer organização. O conceito de protecção de perímetro deixa de ser suficiente assim como o conceito de protecção da rede, passando o foco destes sistemas, nesta terceira geração, a ser a

protecção de todo o negócio. Esta necessidade de olhar para o negócio como um todo surge da evolução que as tecnologias de informação têm vindo a ter. No actual contexto de utilização intensiva das tecnologias fora e dentro do ambiente de trabalho, o grande desafio da segurança está directamente relacionado com as pessoas e os seus comportamentos. Alguns dos principais desafios incluem:

- O número de pessoas constantemente online tem vindo a aumentar e este aumento tem tendência a continuar. As pessoas tendem a aceder e partilhar informação interna e externa à organização, o que faz com que os riscos tendam a aumentar. Estes riscos estão relacionados com o acesso a sites que contêm *software* malicioso e que pode comprometer a máquina de quem está a aceder a estes sites. Por outro lado, a facilidade de acesso à informação pode fazer com que alguns curiosos, várias vezes denominados de hackers principiantes, comecem a tentar qualquer tipo de intrusão ou actividades maliciosas (Nakamura & Geus, 2007);

- O aumento da informação disponível traz, como consequência directa, o aumento do número de equipamentos e o número de pessoas a aceder a estes equipamentos. Estes factos fazem com que exista uma dificuldade acrescida em gerir as actividades de acesso de cada pessoa;

- O número de fusões e aquisições de empresas é também um factor a ter em conta pois resultam em ambientes mais heterogéneos, quer a nível de recursos humanos quer a nível de equipamentos. Com estes ambientes, o número de brechas que podem ser exploradas aumenta o que faz com que haja mais oportunidades para explorar as mesmas. Um bom exemplo deste tipo de brechas é a diferença de configurações em *firewall*s que pertenciam a organizações diferentes. Estas brechas de segurança podem facilmente ser exploradas por *hackers*, aproveitando estes erros de configuração;

- As empresas recorrem cada vez mais a *outsourcing*. Isto implica passar actividades de suporte para o exterior da empresa e, simultaneamente, manter trocas de informação próximas com as empresas de *outsourcing*. Desta forma, aumenta o número de pessoas externas à empresa a trabalhar directamente com informação confidencial e com sistemas internos. As oportunidades de acessos indevidos, quer intencionais quer motivados por comportamentos mais descuidados por parte de pessoas autorizadas, cresceram de forma muito intensa e de difícil controlo (Belcourt, 2006);

- Cada vez mais se recorre ao conceito de *software as a service*. A informação gerida por este *software* fica armazenada nos servidores das empresas que prestam estes serviços, as quais devem igualmente implementar mecanismos de segurança ao mesmo nível das empresas clientes. Caso tal não aconteça, as empresas clientes não conseguem garantir a segurança dessa informação (Choudhary, 2007).

Uma vez que a informação da organização se encontra distribuída por muitos equipamentos, fora e dentro das suas fronteiras, e é acedida diariamente pelos colaboradores e também por pessoas de outras organizações parceiras de negócio, os responsáveis de segurança nas organizações têm de ter cada vez mais em conta as acções das pessoas na rede. Existe a necessidade de saber quem está a qualquer momento na rede, que aplicação é que essas pessoas estão a aceder, que informação possuem ou que informação estão a consultar e que acções tomam com essa informação. Os SIEM da terceira geração vêm precisamente acrescentar a mais-valia de se poder correlacionar informação de utilizadores, equipamentos e rede. Estes sistemas permitem controlar toda a actividade na rede através de *dashboards* personalizados, relatórios, notificações e alertas enviados aos responsáveis. Na figura 11 pode verificar-se um exemplo de dois possíveis ecrãs que se podem obter com as ferramentas de SIEM existentes. É importante verificar que na representação mais à esquerda temos vários tipos de gráfico que podem estar a mostrar a mesma informação de forma diferente ou mesmo diferentes tipos de informação; já na representação mais à direita pode verificar-se uma análise mais pormenorizada sobre determinados eventos.

Figura 11: Exemplo de Ecrã de um SIEM (adoptada de (PunditNetworks, 2011)**)**

Com as tecnologias de SIEM 3.0 existem vários tipos de eventos de diversos equipamentos que podem ser utilizados para correlação, para que se possam detectar comportamentos e actividades que sem esta correlação passariam em claro. Estes tipos de eventos podem ser provenientes de Antivírus, VPN, *firewall*s, *routers, switches*, IDS, IPS, Bases de Dados, Sistemas Operativos, entre outras fontes.

Métricas

Segundo definição do *Systems Security Engineering - Capability Maturity Model* (SSE-CMM), as métricas são medidas quantificáveis de algum aspecto de sistema ou empresa.

Algumas das métricas identificadas para avaliar os resultados de implementações de SIEMs, consoante as várias áreas relacionadas, encontram-se em seguida identificadas e descritas:

Métricas de Carácter Técnico

As métricas de carácter técnico ou tecnológico estão relacionadas com o desempenho de vários processos IT.

- Número de Equipamentos Monitorizados

  A utilização desta métrica pode ser utilizada em várias perspectivas. Os seus resultados fornecem uma visão global da abrangência da monitorização ao longo do tempo. Para além disso pode ser utilizada para efeitos de facturação em ambientes em que o serviço é fornecido por departamentos, empresas do mesmo grupo ou mesmo a clientes.

- Número de Eventos por Equipamento Monitorizado

  A constante monitorização do volume de eventos gerados por cada equipamento ao longo do tempo deve ser uma pratica enraizada. Esta métrica permite uma análise contínua do bom ou mau funcionamento dos equipamentos assim como a identificação de algum comportamento anormal que gere um aumento ou diminuição do numero de eventos gerado pelos por cada equipamento. Esta métrica não tem como objectivo contabilizar eventos correlacionados a menos que o próprio SIEM seja monitorizado pois os próprios eventos internos dos SIEM são também eventos correlacionados.

- Número de Eventos Monitorizados

  Tal como a métrica apresentada anteriormente também esta tem como objectivo analisar o volume de evento com a principal diferença que esta contabiliza todos os eventos incluindo os eventos internos de sistema, eventos *raw* (eventos tal e qual como são gerados nos sistemas) e eventos correlacionados.

- Número de Actualizações por Equipamento Monitorizado

  A utilização desta métrica permite quantificar as actualizações de *software*, assinaturas ou políticas para cada equipamento. Entre outras coisas é possível fazer uma análise do tempo que pode ser dispendido pelos administradores na manutenção dos equipamentos actualizados.

- Número de Indisponibilidades por Equipamento.

  A verificação constante dos resultados desta métrica deve ser encarada como uma boa prática para os profissionais de segurança. Os equipamentos a ser monitorizados, ou mesmo alguns sensores, podem estar com alguma degradação de desempenho que faça com que os serviços fiquem indisponíveis. Esta indisponibilidade pode por em causa a segurança da organização daí a sua importância.

Métricas de Carácter Analítico

As métricas de carácter analítico estão relacionadas com actividades da responsabilidade dos centros de operações de segurança (SOC)

- Número de Sistemas em Conformidade

  Esta métrica é bastante útil não só do ponto de vista de segurança como do ponto de vista de auditoria. Permite identificar o número de máquinas que não estão a cumprir com os requisitos de conformidade estabelecidos.

- Número de Sistemas com *Patchs* Actualizados/Desactualizados

Tal como a métrica anterior, também esta métrica permite identificar o número de máquinas com determinados *patchs* instalados. Para além disso permite o cálculo mais preciso de estimativas de tempo e custo para a manutenção de determinados patchs nas máquinas.

- Número de Sistemas Infectados

  A utilização desta métrica permite não só identificar o número de máquinas infectadas e máquinas não infectadas como permite calcular algumas estimativas sobre o tempo que demora a fazer a desinfecção das mesmas.

- Top de Portas Utilizadas

  A monitorização constante desta métrica assim como os desvios que esta pode apresentar permite identificar não só tendências mas também desvios inesperados que podem ser úteis para a identificação de actividades suspeitas ou anómalas.

- Top de Eventos

  Tal como foi dito para a métrica identificada anteriormente, também esta permite verificar tendências, comportamentos anómalos ou mesmo ataques cujas regras possam não estar a identificar correctamente.

Implementações mais conhecidas

O mercado dos SIEM tem sido nos últimos anos um mercado em evolução. Várias das grandes empresas do mercado das tecnologias de informação como a IBM, a HP, a RSA, entre outras, fizeram também as suas apostas nas soluções de SIEM.

Na figura 12 podemos fazer uma breve análise às principais implementações de SIEM existentes no mercado, assim como uma classificação atribuída pela Gartner consoante algumas das capacidades destes sistemas. Podemos concluir que em duas das principais componentes dos SIEM (Log Management e SEM) o ArcSight da HP é o melhor classificado. Já na componente de relatórios de conformidade as ferramentas da RSA e da SenSage encontram-se melhor posicionadas.

Product Rating	ArcSight/ ESM & Logger	CA/Log Manager	RSA (EMC)/ enVision	IBM/ TSIEM	LogLogic/ Appliances	Novell/ Sentinel and Log Manager	Q1 Labs/ QRadar	SenSage/ Solution	Symantec/ SSIM
Log Management	4.5	2.8	3.6	2.0	4.2	3.2	3.9	4.0	3.6
Compliance Reporting	4.0	3.5	4.4	3.5	3.6	3.0	3.8	4.5	2.8
SEM	4.9	1.0	2.7	2.8	2.2	4.3	4.1	2.1	3.8
User Monitoring	4.6	3.8	3.5	3.7	2.7	3.9	3.5	4.0	2.7
Application Monitoring	4.8	2.5	3.3	2.3	3.5	3.2	3.3	4.3	3.4
Deployment and Support Simplicity	2.8	2.5	4.5	2.0	3.5	2.5	4.3	2.5	3.5

Figura 12: Classificação dos SIEM consoante as suas capacidades (0-5) (adaptada de (Nicolett, 2010)**)**

Capitulo IV – Método de Implementação de SIEMs

Com o método de implementação de SIEM proposto neste capítulo, pretende-se apresentar um modelo para a acção do implementador bem como apoio nas decisões que terá que tomar ao longo de todo o processo de implementação. Este método é constituído por 5 fases: planeamento, instalação, recolha de eventos, optimização, e operação e administração. Cada fase é constituída por várias actividades que o implementador deverá realizar. A figura 13 apresenta uma visão global do método e as secções seguintes descrevem em detalhe as suas várias fases. O resultado aqui apresentado foi sendo construído com base na experiência que o investigador adquiriu, como colaborador da empresa Unisys, na implementação de diferentes projectos de SIEM. A implementação destes projectos, inicialmente acompanhado por profissionais mais experientes, permitiu que fossem identificados alguns pontos em que normalmente fazem com que os projectos possam ter resultados inesperados sendo que estes pontos são mencionados ao longo desta secção.

De realçar que apesar de, por razões de simplicidade de apresentação, as fases e actividades serem apresentadas de forma sequencial, reconhece-se que a realização de uma actividade pode levar à reformulação de uma actividade anterior. No entanto, dado que algumas das decisões se tornam difíceis de alterar posteriormente, também se reconhece que o implementador se deve munir de toda informação relevante para minimizar a necessidade de refazer actividades anteriores.

Figura 13: Método de Implementação de SIEMs

Planeamento

A fase de planeamento da implementação de SIEMs é a fase mais importante do ciclo de vida e também a fase mais morosa de um projecto deste tipo. Uma grande parte dos projectos fica aquém do esperado devido à pouca atenção que normalmente é dedicada a esta fase.

Ao longo desta secção são apresentados e explicados os aspectos que se deve ter em conta para que o sucesso do projecto não seja comprometido. Tarefas relacionadas com a definição de objectivos, definição do âmbito, definição dos casos de uso, eventos, arquitectura e estrutura de responsabilidades. Esta é uma sequência que permite identificar alguns limites do projecto, tal como os equipamentos que precisam ser monitorizados, o tipo de informação a analisar, os recursos humanos envolvidos e os requisitos de *hardware*/ *software* necessários.

Definir Objectivos

A primeira tarefa do planeamento é a tarefa da definição de objectivos. A decisão pela implementação de um SIEM é consequência da identificação de necessidades específicas de gestão da segurança. Para se chegar à conclusão de que há necessidade de implementar um SIEM é preciso que já se faça alguma monitorização da actividade nos sistemas de forma a saber o que se pretende do SIEM: fazer gestão de eventos (normalmente de forma centralizada); monitorizar aspectos relacionados com a segurança; resolver problemas conhecidos de conformidade; ou então a combinação entre qualquer destas necessidades.

As empresas que optam por SIEMs possuem, habitualmente, outro tipo de sistemas como IDSs, IPSs, *firewalls*, antivírus, entre outros. No entanto, individualmente ou em conjunto, estes sistemas não têm a capacidade de gestão centralizada e normalizada de eventos nem a capacidade de fazer a correlação dos mesmos, sendo estes provenientes de diversos sistemas. Um evento por si só pode não ser indício de qualquer actividade maliciosa. No entanto, correlacionado com outros eventos, pode ser o suficiente para a detecção de um ataque. Estes três aspectos são os mais importantes a ter em conta na definição de objectivos para um SIEM: centralização, normalização e correlação.

Definir o Âmbito

A definição do âmbito da implementação de SIEM é uma das tarefas mais importantes uma vez que o resultado desta tem um grande impacto em várias questões do projecto, como por exemplo o investimento financeiro necessário, dependências, requisitos, o tempo de vida do

projecto, as pessoas envolvidas, os riscos associados, oportunidades, pressupostos e restrições (Miguel, 2006).

Dependendo do tipo de negócio em questão, o âmbito pode ser muito variado. Um dos aspectos a ter inicialmente em causa para limitar o âmbito deve ser a própria Organização e decidir se o SIEM a vai abranger totalmente, apenas a parte interna da rede, apenas a parte externa da rede (perímetro), apenas os colaboradores, apenas os clientes ou ambos. Após esta decisão estar tomada, há outra importante a tomar, que é decidir se o âmbito se restringe apenas à infra-estrutura da organização, apenas à informação ou ao conjunto dos dois.

Definir *Use Cases* e Equipamentos a Monitorizar

No contexto dos SIEM, um *Use Case* ou caso de uso é a forma como determinada informação, relacionada com questões de segurança ou requisitos de negócio, é apresentada e visualizada.

Os *Use Case* são definidos de acordo com os objectivos e o âmbito do projecto. Nesta tarefa são definidos inicialmente os *Use Case* genéricos e quais os equipamentos da infra-estrutura que fornecem o tipo de informação necessária para a concepção dos mesmos. Com base nestes *Use Case* vão sendo identificados *Use Case* mais específicos que permitam tirar vários tipos de conclusões.

De seguida são apresentados alguns dos *Use Case* mais genéricos, que normalmente são identificados como importantes do ponto de vista da segurança. Estes *Use Case* genéricos estão directamente associados a diferentes tipos de tecnologias:

Relatórios "inter-dispositivos"

- Top de utilização de largura de banda por utilizador
- Alterações de configurações
- Logins falhados e bem sucedidos
- Alterações de *passwords*
- Top de origens e destinos

Relatórios de Antivírus

- Top de sistemas infectados
- Todos os Erros de AV
- Estatística de actualização de assinaturas de AV
- Actividade AV consolidada
- Alterações de configuração AV

Bases de dados

- Logins falhados e com sucesso
- Alterações de configuração

IPS/IDS

- Top de origens e destinos de alertas
- Top de origens e destinos

Gestão de acessos

- Autenticações com sucesso e falhadas

- Alterações de configuração de gestão de utilizadores

Dispositivos de rede

- Utilização de largura de banda
- Alterações de configuração por utilizador e tipo de alteração
- Logins falhados e bem sucedidos
- Top de ligações

Relatórios de dispositivos VPN

- Número de ligações
- Duração de ligações
- Ligações aceites e rejeitadas
- Logins falhados e com sucesso
- Top de ligações
- Top de utilizadores por utilização de largura de banda
- Alterações de configurações VPN

Sistemas operativos

- Administração de utilizadores privilegiados
- Logins falhados e com sucesso
- Alterações de configurações

Firewall

- Ligações rejeitadas (do exterior)
- Ligações rejeitadas (para o exterior)
- Utilização de largura de banda
- Logins falhados e com sucesso

A forma como estes casos de uso são visualizados pode variar. Alguns destes casos de uso requerem uma análise mais profunda e nesse aspecto devem ser utilizados relatórios. Já quando apenas há interesse em visualizar constantemente por exemplo o top 10 ou 20 de alguma actividade, os *dashboards* são um recurso mais indicado.

Em seguida podemos verificar casos de uso mais específicos que fazem mais sentido visualizar em relatórios e em dashboards:

Relatórios

- Alterações de configurações de antivírus, por utilizador;
- Alterações de configurações de antivírus, por tipo de configuração;
- Largura de banda utilizada, por protocolo;
- Logins falhados por endereço de destino;
- Logins com sucesso por endereço de destino;
- Utilizadores com palavras-chave alteradas;
- Alertas gerados pelos IDSs
- Ligações VPN em período nocturno;
- Autenticações falhadas na VPN

Dashboards

- Top 10 de vírus detectados;
- Top 10 de erros de antivírus;
- Top 20 de acessos negados a determinada porta;
- Top 15 de Utilizadores com logins falhados;
- Top de utilizadores por número de ligações;
- Top 10 de alertas dos IDSs
- Top 20 Países com ligações à VPN

Identificar Tipos de Eventos a Monitorizar

Após identificados os casos de uso e os equipamentos a monitorizar, existe a necessidade de verificar que tipos de eventos se pretendem recolher para se efectuar qualquer tipo de análise.

Segundo recomendações do NIST (agência federativa de tecnologia que funciona com a industria para desenvolver e aplicar tecnologias, medidas e *standards*) na publicação especial 800-92, existem algumas perguntas que têm de ser respondidas (Kent & Souppaya, 2006):

- Que tipos de equipamentos devem ou deveriam registar eventos?
- Que tipos de componentes dos equipamentos devem ou deveriam registar eventos? (ex: sistema operativo, aplicações)
- Que tipos de eventos cada componente deve registar? (ex: eventos de segurança, eventos de conectividade de rede)
- Que tipo de informação deve conter cada tipo de eventos? (ex: endereço destino, nome do utilizador)
- Com que frequência cada tipo de eventos deve ser registado?

Dos equipamentos pertencentes à infra-estrutura da organização, todos ou quase todos geram diferentes tipos de eventos de acordo com os tipos de actividade de cada um. O que acontece é que nem todos os eventos gerados por estes equipamentos têm utilidade significativa no que respeita aos casos de uso que se pretendem alcançar e às actividades que se pretendem monitorizar.

Temos como exemplo os "*domain controlers*" que geram vários tipos de eventos entre os quais temos os eventos relacionados com autenticação, os eventos relacionados com acesso a ficheiros, entre outros. Nesta tarefa há necessidade de investigar todos os tipos de eventos e definir os que estão dentro do âmbito de forma a eliminar informação inútil que pode degradar a

performance dos SIEMs desnecessariamente. Se for dada atenção aos exemplos de casos de uso identificados na tarefa anterior, podemos facilmente concluir que os eventos relativos a acções de *login* são eventos necessários; já os eventos relativos a *logout* ou eventos de acessos a ficheiros não têm qualquer importância para satisfazer os casos de uso.

Os eventos têm associados números de identificação mediante o tipo de evento. Por exemplo, os eventos de *login* com sucesso nos sistemas Windows tem como identificador (*event* ID) o número 528, já os eventos de *login* com sucesso na rede têm associado o identificador 540. Na tabela 6 são apresentados alguns sites onde pode ser consultada informação relativa a eventos, quer sobre o seu formato quer sobre a informação que os eventos contêm.

Log Type	URL
Firewall logging and monitoring	*http://www.loganalysis.org/sections/parsing/application-specific/firewall-logging.html*
Linux system log management and monitoring	*http://www.oreilly.com/catalog/bssrvrlnx/chapter/ch10.pdf* (excerpt of Building Secure Servers with LINUX by Michael D. Bauer)
Microsoft log events (Events and Errors Message Center)	*http://www.microsoft.com/technet/support/ee/ee_advanced.aspx*
Microsoft Windows 2000 logs	Chapter 9, "Auditing and Intrusion Detection", of Securing Windows 2000 Server, *http://www.microsoft.com/technet/security/prodtech/windows2000/secwin2k/default.mspx*
Microsoft Windows Security Log Encyclopedia	*http://www.ultimatewindowssecurity.com/encyclopedia.html*
Microsoft Windows Server 2003 logs	*http://www.microsoft.com/technet/security/prodtech/windowsserver2003/w2003hg/sgch00.mspx*
Microsoft Windows log management script	*http://support.microsoft.com/?id=318763*
Microsoft Windows XP event log management	*http://support.microsoft.com/?scid=308427*
Web server common log file format	*http://www.w3.org/Daemon/User/Config/Logging.html*

Tabela 5: Informação e formatos comuns de eventos (adoptada de (Kent & Souppaya, 2006)**)**

Há determinados modelos de negócio que têm requisitos de conformidade que é necessário respeitar e nesse aspecto a identificação do tipo de eventos a monitorizar tem de ter este aspecto em conta. Exemplo disso são as organizações cujo negócio passa pelos pagamentos electrónicos pois estas têm de respeitar a norma PCI DSS (*Payment Card Industry Data Security Standard*) que exige que todos os acessos a recursos na rede e acessos a informação de portadores de cartões têm de ser monitorizados *("Payment Card Industry Data Security Standard,"* 2009).

A tarefa de investigar todos os tipos de eventos de todos os equipamentos é uma tarefa que requer algum esforço, quer da parte de quem adquire as tecnologias de SIEM quer de quem as implementa. Uma boa prática de optimizar este processo é envolver nesta tarefa os administradores de cada equipamento pois estes têm normalmente algum conhecimento no que diz respeito ao conteúdo dos eventos gerados e à sua utilidade.

Calcular o Número de Eventos por Segundo

Uma das tarefas importantes da fase de planeamento é a recolha e informação relativa ao número de eventos que cada equipamento a monitorizar gera por segundo. Este tipo de informação é muito importante para o dimensionamento da solução implementada quer a nível da arquitectura quer para a análise dos requisitos de *hardware* que a plataforma SIEM necessita.

O número de eventos gerados em determinada organização está directamente relacionado com a sua politica de segurança, isto é, nem todas as acções que ocorrem na rede necessitam de ser registadas como um evento e essa decisão deve estar reflectida na política de segurança da organização em causa. Também a incorrecta configuração dos equipamentos pode fazer com que, por exemplo uma firewall possa registar eventos num número muito superior ao que é necessário (Kent & Souppaya, 2006).

Para que o número de eventos por segundo seja um valor o mais realista possível, o ideal é recolher informações ao longo de um mês completo ou, na pior das hipóteses, ao longo de uma semana. Isto deve-se ao facto de o número de eventos gerado pelos equipamentos ter oscilações ao longo do dia e mesmo ao longo da semana, como por exemplo durante a noite ou ao final de semana, respectivamente. Após esta recolha basta fazer a razão entre o número total de eventos recolhidos e o período de tempo que efectivamente se contabilizou a recolha de eventos. Nesta altura, se a organização já tiver alguma tecnologia de centralização de eventos, o cálculo do número de eventos por segundo fica mais facilitado, caso contrário poderá ser necessário fazer a contagem dos eventos equipamento a equipamento (Angelino, n d).

Uma vez que esta tarefa pode trazer algumas dificuldades e por vezes não seja fácil contabilizar o número de eventos por segundo, na figura 14 podemos verificar alguns valores de referência para sistemas utilizados nas infra-estruturas das empresas. Estes valores têm como base 750 utilizadores, 750 *endpoints*, 5 localizações, 6 sub-redes, 5 bases de dados e um centro de dados.

Qty	Type	Description	Avg EPS	Total Peak EPS	Average Peak EPS
750	Employees/Endpoints (Windows XP)	Desktops & laptops at 5 locations	Included at domain servers	Included at domain servers	Included at domain servers
7	Cisco Catalyst Switches	One at each location, one in DMZ and one in the Trusted network	5.09	51.88	26.35
7	Cisco Gateway/Routers	One at each location	0.60	380.50	154.20
5	Windows 2003 Domain Servers	One at each location	40.00	404.38	121.75
3	Windows 2003 Application Servers	In high availability cluster at data center	1.38	460.14	230.07
3	MS SQL Database Servers running on Windows 2003 Server	High availability cluster at data center	1.83	654.90	327.45
6	Microsoft Exchange Servers	One at each location with two (cluster) at the data center	3.24	1,121.50	448.60
3	MS IIS Web Servers on Windows 2003	High availability cluster at data center	1.17	2,235.10	1,117.55
2	Windows DNS Servers	At data center – failover	0.72	110.80	110.80
2	Linux Legacy Application Servers	At data center	0.12	43.60	21.80
1	Linux MySQL Database Server	One in Trusted network for legacy application	0.12	21.80	21.80
7	NitroGuard IPS	One at each location, one in DMZ and one in the Trusted network	40.53	5,627.82	1,607.95
1	Netscreen Firewall	Netscreen facing the Internet	0.58	2,414.00	2,414.00
3	Cisco Pix Firewalls	Between the data center and the other four sites, in front of Trusted network, between Trusted and the DMZ	39.00	1,734.00	1,178.00
1	Cisco VPN Concentrator	Located at data center Facing the Internet	0.83	69.45	69.45
1	Squid Proxy	Located at data center	14.58	269.03	269.03
		Totals:	**149.79**	**15,598.90**	**8,118.80**

Figura 14: Valores de referência de EPS para equipamentos de rede (adoptado de (Butler, 2009)**)**

Definir os Períodos de Retenção

Uma das capacidades que os SIEMs têm é a de armazenamento de eventos. Como a capacidade dos sistemas de armazenamento não é ilimitada, é pertinente definir por quanto tempo devem ser retidos os eventos. Este período de armazenamento de eventos é definido com base em dois requisitos chave. Um deles é o requisito legal que, para cada organização, e mediante o seu negócio, pode apresentar um período legal muito diferente. Por exemplo, a norma *PCI DSS* menciona um período de retenção de eventos de pelo menos um ano, sendo

que pelo menos três meses têm de estar disponíveis online. Normalmente este requisito legal está reflectido na política de segurança da organização. Portanto, se esta politica estiver actualizada e de acordo com os requisitos legais não há necessidade de voltar a consultar todas as normas na altura em que se define o período de retenção dos eventos. O segundo critério está relacionado directamente com os sistemas cujos eventos estão a ser armazenados, mais concretamente com a classificação que estes têm associada relativamente ao impacto que estes sistemas podem causar.

Na tabela 7 podem ser analisadas algumas definições de configuração a aplicar sobre os componentes de gestão dos eventos.

Category	Low Impact Systems	Moderate Impact Systems	High Impact Systems
How long to retain log data	1 to 2 weeks	1 to 3 months	3 to 12 months
How often to rotate logs	Optional (if performed, at least every week or every 25 MB)	Every 6 to 24 hours, or every 2 to 5 MB	Every 15 to 60 minutes, or every 0.5 to 1.0 MB
If the organization requires the system to transfer log data to the log management infrastructure, how frequently that should be done	Every 3 to 24 hours	Every 15 to 60 minutes	At least every 5 minutes
How often log data needs to be analyzed locally (through automated or manual means)	Every 1 to 7 days	Every 12 to 24 hours	At least 6 times a day
Whether log file integrity checking needs to be performed for rotated logs	Optional	Yes	Yes
Whether log data transfers to the log management infrastructure need to be encrypted or performed on a separate logging network	Optional	Yes, if feasible	Yes

Tabela 6: Exemplo de definições de configuração para a gestão de eventos (adoptado de (Kent & Souppaya, 2006)**)**

Para a tarefa em questão é importante reflectir acerca da primeira e da segunda entrada da tabela acima representada, onde é indicado um acréscimo do tempo do período de retenção à

medida que o impacto aumenta. Para além disso, verifica-se que quanto maior o impacto dos sistemas, menor deve ser o tempo de rotatividade dos eventos.

Definir Arquitectura Física

A definição da arquitectura física dos sistemas envolvidos no projecto é também uma das tarefas importantes da fase de planeamento.

Praticamente todas as tecnologias de SIEM das várias empresas têm disponíveis *appliances* pré-instaladas. Na altura em que se planeia a arquitectura há algumas decisões que têm de ser tomadas, sendo que uma delas é a escolha entre *appliances* pré-instaladas e servidores para suportar as instalações do *software*. Nesta altura é necessário também ter em conta que as *appliances* têm várias limitações relacionadas com o processamento de EPS, o que pode excluir imediatamente esta opção.

Uma vez seleccionado o *hardware* que irá suportar o *software* dos SIEMs, é necessário identificar como serão extraídos os eventos dos sistemas de origem. Há três questões que devem ser tidas em causa para garantir que não falham pormenores importantes:

- Como é que os eventos são transferidos da fonte para o SIEM? (ex: protocolo utilizado)
- Com que frequência é que os eventos devem ser transferidos dos equipamentos para a plataforma de SIEM? (ex: em tempo real, ao fim do dia)
- Como é que a confidencialidade, integridade e disponibilidade de cada tipo de evento deve ser garantida enquanto estes transitam dos sistemas de origem para a plataforma SIEM? (ex: encriptados)

Esta identificação tem de ser efectuada caso a caso uma vez que em determinadas situações pode ser necessária a utilização de *software* de terceiros. Exemplos disso são os sistemas de *mainframe* que normalmente registam os eventos em formato binário. Nestes casos é necessário um *software* intermediário que processe esse ficheiro binário e produza um ficheiro legível aos sistemas de SIEM, para que a normalização possa ser feita de forma correcta.

A figura 15 apresenta uma arquitectura de uma plataforma SIEM que, neste caso concreto, utiliza a ferramenta de SIEM da ArcSight. Analisando a figura de baixo para cima, podemos verificar na camada mais inferior que estão a ser monitorizados equipamentos de 3 locais distintos, sendo que nesses 3 locais existem diferentes conectores que processam os eventos e os encaminham para a componente de gestão de eventos proveniente dos sistemas SEM, de acordo com a sua localização (segunda camada). Com as setas verdes existentes neste segunda camada podemos concluir que as componentes de gestão de eventos dos 3 locais distintos

estão integradas entre si, o que permite uma consulta dos eventos a partir de um único ponto. Após os eventos chegarem aos componentes de gestão de eventos todos eles são armazenados e alguns deles são encaminhados para a componente de correlação do respectivo local, para serem correlacionados em tempo real (terceira camada). Os eventos que resultam do processo de correlação são também enviados para a componente de gestão de eventos mas, para além disso, os eventos correlacionados resultantes da correlação por localização são reencaminhados para um motor de correlação global (quarta camada) que faz a correlação dos eventos provenientes dos vários locais. Verifica-se na imagem que, em alguns casos, existe mais do que um componente por local, o que representa um servidor adicional que pode ser para alta disponibilidade ou simplesmente para suportar uma base de dados. No canto superior esquerdo da figura estão representados os analistas que, através de uma consola, realizam o seu trabalho com base em eventos armazenados na componente de gestão de eventos, com base nos eventos correlacionados de cada local e com base nos eventos correlacionados resultantes da correlação entre os vários locais.

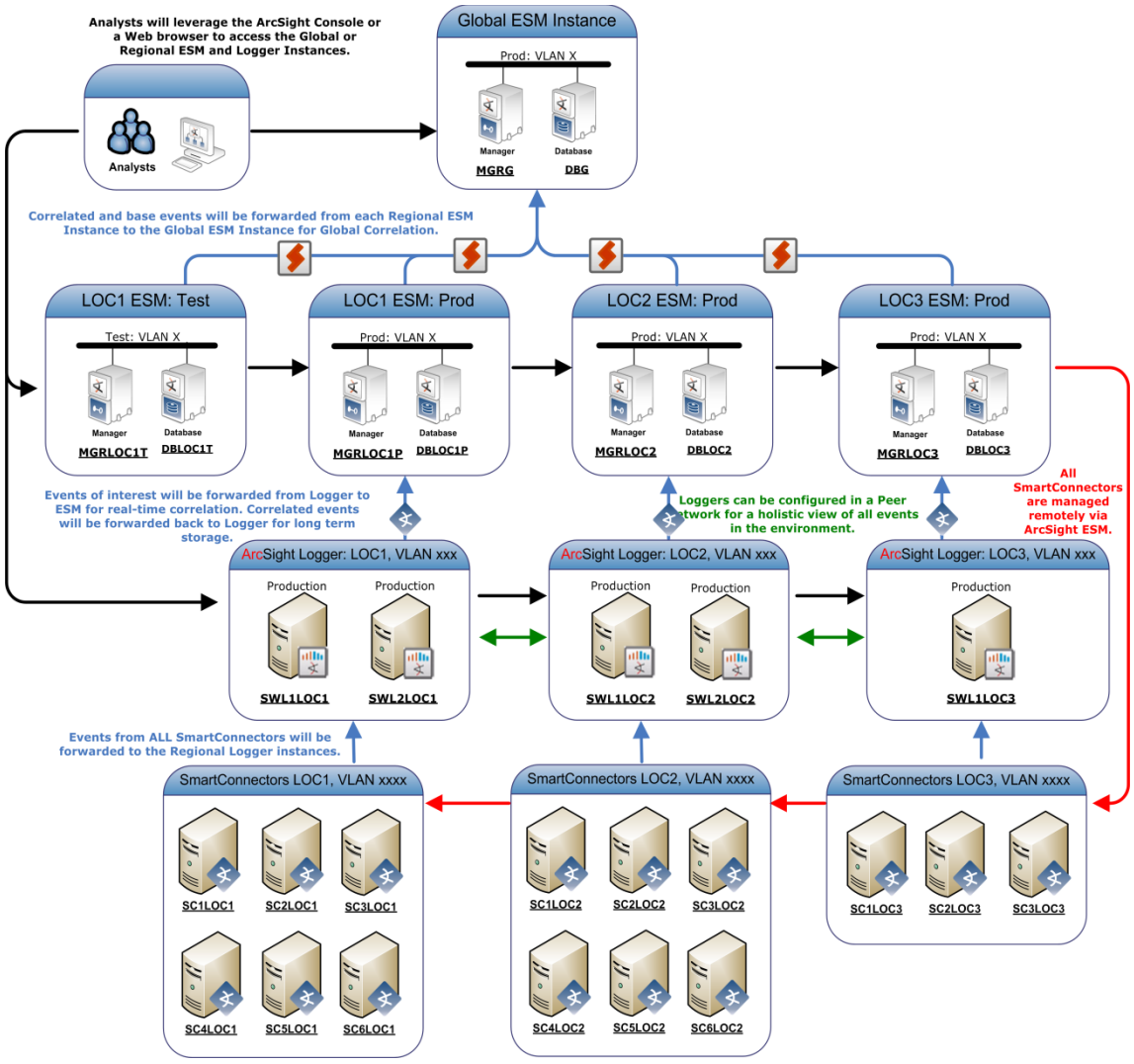


Figura 15: Exemplo de Arquitectura com ArcSight SIEM

Identificar Perfis e Utilizadores

Um dos aspectos a ter em conta na fase de planeamento é o tipo de perfis e quais os utilizadores que irão interagir com a plataforma. Apesar de estes perfis e utilizadores poderem vir a sofrer alterações no futuro, inicialmente é necessário proceder com esta definição.

Do facto de os SIEMs permitirem a consulta da informação proveniente de várias origens, surge a necessidade de restringir o acesso a essa informação. Por exemplo, os colaboradores responsáveis pelas comunicações apenas devem poder consultar informação relacionada com comunicações, os responsáveis pelas redes apenas devem poder consultar informação relacionada com redes, etc.

Para além disso, existem normalmente outros dois tipos de perfil que são criados: Administradores e Auditores. Os administradores são normalmente um número reduzido de

pessoas que têm todo o tipo de privilégios sobre o SIEM, desde a consulta de informação, ao desenvolvimento de conteúdos (por exemplo relatórios) ou mesmo à alteração de configurações. Já os perfis de auditor, tal como o nome indica, apenas permitem consultar a informação contida na plataforma. Normalmente estes auditores são colaboradores internos e não auditores externos.

Definir Hierarquia de Responsabilidades

Após definidos os perfis e respectivos utilizadores, é necessário definir uma hierarquia de responsabilidades. Esta hierarquia de responsabilidades é de grande utilidade nos alertas e nas notificações que são gerados pelos SIEMs.

Quando se está implementar um destes sistemas, é importante estar definido quem vai responder por determinados alertas ou notificações. As notificações relacionadas com o *hardware* destes SIEMs devem ser uma prioridade dado que estes sistemas passam a conter informação centralizada de vários equipamentos. Discos, ventoinhas, processadores e memórias são alguns dos recursos que devem ser alvo de uma monitorização constante para que se possam minimizar as perdas de informação.

Instalação

Uma vez terminada a fase de planeamento, onde se tomaram decisões importantes para todo o projecto, passa-se então para a fase de Instalação. Esta fase de instalação é composta por três tarefas, nomeadamente a Recolha e Aplicação de Informação, a Instalação de *Patchs* e os Testes de Conectividade.

Recolher e Aplicar Informação

Uma vez na fase de instalação é necessário começar por recolher alguma informação exigida pelos SIEMs como por exemplo, os *hostnames* que são os nomes pelo qual cada equipamento vai ser conhecido na rede, os IPs e respectiva máscara, a *gateway*, as caixas de correio electrónico, entre outra que pode ser relacionada com a organização em causa.

Para além desta informação, é necessário recolher e aplicar informação relativa a alguns serviços:

- IP do Serviço NTP: é importante que todos os sistemas estejam sincronizados com um servidor de NTP para que estes partilhem sempre da mesma hora e quando possível, deve-se indicar o IP de um servidor de NTP secundário para o caso de falha do servidor principal.

- IP do Serviço de DNS: os serviços de DNS irão permitir que seja feita a resolução de nomes de domínios em endereços IP. Esta é uma funcionalidade fundamental para quase todos os SIEMs uma vez que, em alguns eventos apenas está registado o nome do domínio. No processo de normalização os SIEMs resolvem estes nomes para os respectivos endereços da rede.
- IP do Serviço SMTP: os serviços SMTP permitem o envio de emails através da rede. É importante definir estes servidores devido à possibilidade de envio de notificações por correio electrónico, que naturalmente irá utilizar estes servidores.
- IP do Serviço de *Active Directory* ou do Serviço de RADIUS: a autenticação integrada com um destes servidores é extremamente importante, uma vez que limita bastante o acesso à plataforma por pessoas indevidas. Quando esta integração não é feita há necessidade de criar contas de login locais o que pode ser visto como uma vulnerabilidade relevante.

Instalar *Patchs*

Após definida a arquitectura física para o SIEM, já é conhecido o tipo de equipamentos que serão utilizados e com base nisso, a instalação de *patchs* vai ser feita sobre as *appliances* adquiridas e/ou sobre o os equipamentos disponibilizados para a instalação do *software*.

É muito importante que os equipamentos escolhidos sejam integrados na rede já com as ultimas versões e *patchs* instalados. Apesar de estas instalações poderem ser feitas posteriormente, é importante que seja feito nesta fase uma vez que estes *patchs* corrigem normalmente erros ou vulnerabilidades que podem por em causa o bom funcionamento dos sistemas.

Nos tempos de hoje a gestão de *patchs* é uma questão crítica pois passou a ter uma grande importância para garantir o bom funcionamento e a salvaguarda da informação de qualquer organização. São conhecidos incidentes no passado em que organizações foram atacadas com base em vulnerabilidades conhecidas e o ataque terminou com sucesso uma vez que essas organizações não tinham instalado os *patchs* que corrigiam a situação (Cavusoglu & Zhang, 2008).

Efectuar Testes de Conectividade

Uma vez que todos os equipamentos têm toda a informação necessária aplicada e estão com os últimos *patchs* instalados, passa-se para a tarefa de realizar testes de conectividade, entre eles e com os servidores que foram indicados anteriormente (NTP, DNS, SMTP).

É perfeitamente natural que possam haver *firewalls* entre estes sistemas e os servidores que se pretende testar a conectividade. Portanto têm de ser criadas excepções nessas mesmas *firewalls* de modo a que estas permitam que estas máquinas comuniquem entre si.

Recolha de Eventos

Instalar Conectores/Colectores

Os conectores, por vezes também conhecidos por colectores, são os componentes existentes nos SIEM responsáveis por fazer a processamento dos eventos provenientes dos diversos sistemas. Dependendo do tipo de equipamento e das opções tomadas na definição da arquitectura, os eventos podem ser extraídos pelos conectores directamente da sua origem ou então podem ser enviados da origem para um servidor que faz o processamento dos mesmos. Por exemplo, alguns equipamentos têm uma opção que permite enviar os eventos através do protocolo de *syslog* para um servidor que tem um conector à espera de receber esses eventos.

Para além de fazerem a recolha dos eventos, na grande maioria dos SIEM são os conectores que têm a função de os normalizar para que estes possam ser geridos todos da mesma forma e para facilitar o trabalho no motor de correlação. É comum existirem diferentes tipos de conectores para diferentes tipos de equipamentos. Estes conectores estão parametrizados para processar diferentes tipos de eventos, dependendo da sua origem.

Na grande maioria dos SIEM existem várias formas de gerir os conectores sendo que três delas são as mais aconselhadas. Uma das hipóteses é recorrer a *appliances* que contém já os conectores previamente instalados e apenas é necessário proceder à sua configuração. Outra solução é instalar os conectores e configurá-los em servidores dedicados. A terceira opção é instalar os conectores nas próprias máquinas onde os eventos são gerados. A escolha por qualquer uma destas opções é feita normalmente com base em dois factores: o preço e a politica da organização para o tipo de projecto em causa, tendo em conta que não há limitações por parte de qualquer equipamento.

Nesta tarefa decorrem todas estas decisões e prossegue-se com a configuração dos conectores necessários para centralizar todos os eventos dos diferentes equipamentos, previamente definidos para o âmbito do projecto.

Configurar os Equipamentos a ser Monitorizados

A configuração dos equipamentos que serão monitorizados pode ter de ser feita depois de os conectores estarem configurados ou ao mesmo tempo. Raros são os casos em que a configuração dos equipamentos deve ser feita em primeiro lugar.

Este tipo de configurações deve ser efectuado pelo administrador do equipamento em causa uma vez que este deve ser a pessoa com mais conhecimento sobre o mesmo.

Existem vários tipos de configuração nos equipamentos que podem ter de ser feitos e que podem ir desde a activação de uma opção de *logging* até ao desenvolvimento de scripts. Por exemplo, a recolha de eventos das *firewall CheckPoint* requer na maior parte dos casos o estabelecimento de uma conexão, recorrendo ao protocolo OPSEC, o que obriga a que a configuração tenha de ser feita com interacção entre ambas as partes e em paralelo; já na maior parte das versões das *firewalls Juniper* os eventos são enviados pelos equipamentos recorrendo ao protocolo *syslog* e, neste caso, esta configuração pode ser efectuada a qualquer altura, independentemente do processo de instalação do conector.

Na figura 16 podemos verificar o painel de propriedades sobre os eventos de segurança do Windows. Neste caso, os principais aspectos a ter em conta são o local onde os eventos serão guardados, qual o tamanho do ficheiro que irá registar os eventos e para além disso confirmar que a opção de registo de eventos está activa.

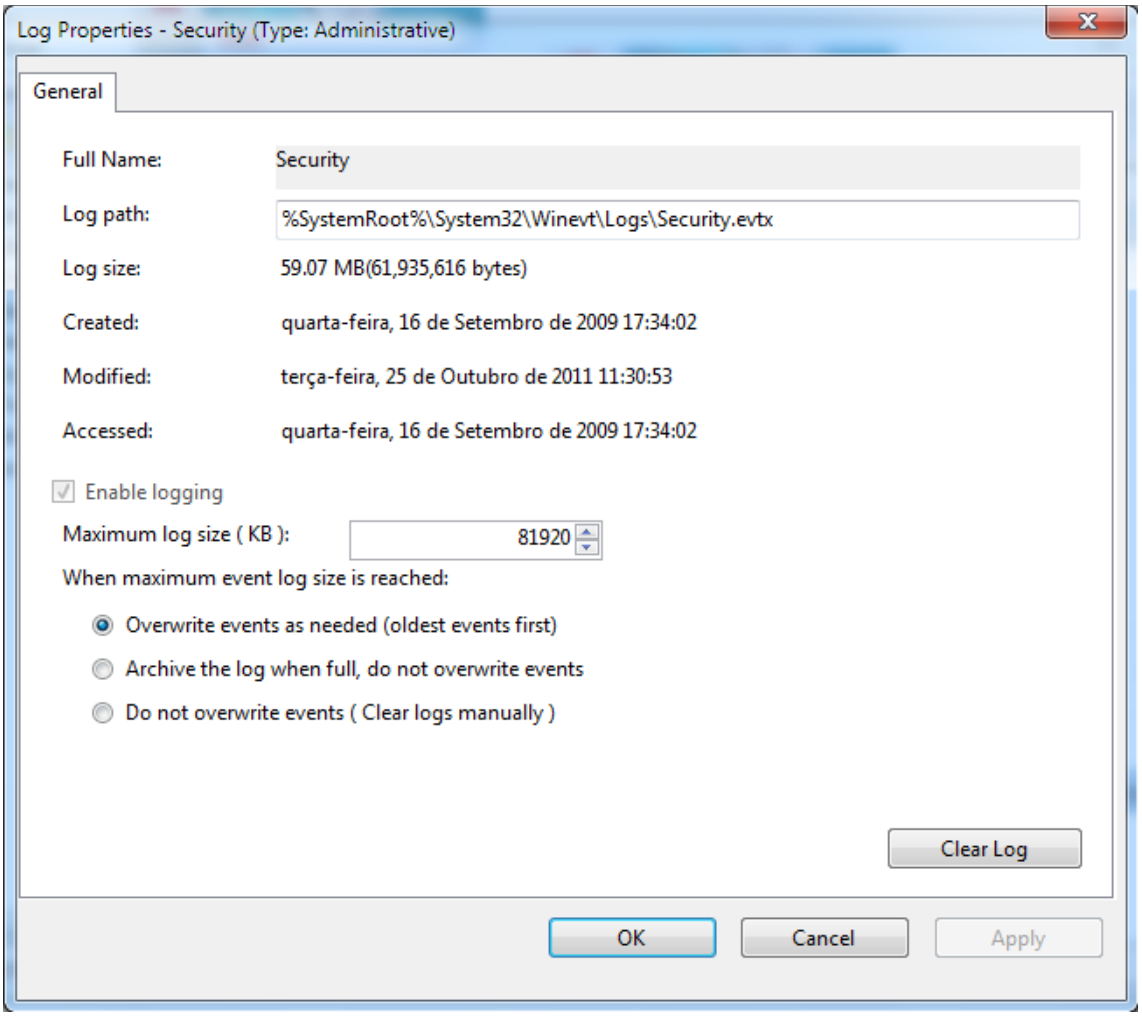


Figura 16: Configuração dos eventos de segurança do Windows 7

Validar a Normalização dos Eventos

Uma vez que os eventos estão a chegar à plataforma de SIEM é necessário proceder a uma verificação pormenorizada dos mesmos para verificar se a normalização dos mesmos ocorreu consoante esperado. Esta normalização é extremamente importante uma vez que é com base em eventos normalizados que serão feitas as pesquisas, os relatórios, os filtros, as regras de correlação e todas as outras operações possíveis.

É no processo de normalização que os eventos são categorizados e que lhes é atribuída uma classificação de severidade que pode não ser a mais adequada, dependendo do ambiente. Nesta tarefa estes dois aspectos não devem passar despercebidos uma vez que são características importantes para a administração e operação de toda a informação dentro da plataforma.

Uma vez detectadas anomalias provenientes do processo de normalização é necessário alterar as configurações de mapeamento ou categorização, na maioria dos casos dos conectores uma vez que são os responsáveis pelo processo.

Desenvolver Conectores à Medida

Os SIEMs são compostos por uma série de conectores *out-of-the-box* que já estão desenvolvidos e fazem já o mapeamento dos eventos para a estrutura do SIEM. Normalmente estes conectores são desenvolvidos para equipamentos mais utilizados no mercado. No entanto, para os equipamentos em que não há conectores previamente desenvolvidos, há necessidade de os desenvolver à medida.

Por exemplo, os fornecedores de SIEM não desenvolvem conectores para aplicações desenvolvidas internamente nas empresas. Portanto, nestas situações, há necessidade de desenvolvimento de conectores à medida, que façam o registo dos eventos. Este detalhe pode revelar-se um problema de difícil resolução, em especial com aplicações antigas, relativamente às quais os próprios fornecedores já perderam alguma capacidade de desenvolvimento.

Instalar Agentes Intermediários

Por uma questão de optimização ou por uma questão de necessidade, por vezes pode ser necessário instalar agentes intermediários ou criar *scripts* que façam algum tipo de actividade no processo de passagem dos eventos entre os equipamentos e os conectores. Um bom exemplo em que normalmente este problema surge é na recolha dos registos do sistema operativo z/OS. Estes eventos vão sendo registados num ficheiro, em binário, existindo agentes que processam esses ficheiros e que geram como *output* um ficheiro legível em ASCII, sendo esse o ficheiro que é processado posteriormente pelos conectores dos SIEM.

Um outro agente bastante utilizado designa-se por *Snare* da *IntersectAlliance*. Os eventos do Windows normalmente têm um tamanho maior do que eventos de outros sistemas. Assim, em várias implementações é comum instalar-se este ou outros agentes que recolhem os eventos do Windows e os transformem eventos adequados para um servidor de *syslog.*

Optimização

Ajustar Configuração dos Conectores

Após os conectores estarem instalados e a plataforma de SIEM estar a receber eventos, convém que seja feita uma análise pormenorizada, conector a conector, para que seja encontrada uma configuração próxima do ideal, para cada caso. Os SIEM permitem a filtragem e agregação de eventos à entrada para a plataforma e estes dois parâmetros requerem bastante atenção. Para o processo de filtragem de eventos é necessário analisar cada tipo de eventos e verificar se há interesse que este seja mantido na plataforma de SIEM ou não. Já para o processo de

agregação, é necessário ter em atenção quais os campos que são identificados e a partir dos quais se agregam os eventos, pois uma má identificação destes campos pode levar a que algumas regras de correlação fiquem obsoletas. Para além destes dois parâmetros principais há outros que devem ser tidos em conta, tal como a capacidade de cache de cada conector (necessária para o caso de a ligação falhar), a utilização da largura de banda a ocupar, entre outros parâmetros que variam de fabricante para fabricante.

Nas figuras 17 e 18, retiradas directamente de *ArcSight* SIEM, podemos verificar alguns exemplos dos parâmetros importantes que devem ser configurados com algum cuidado:

Figura 17: Configuração do parâmetro *Field Based Agregation*

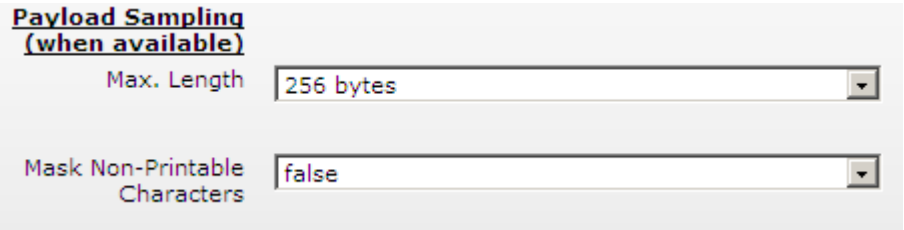


Figura 18: Configuração do Parâmetro *Payload Sampling*

Aplicar Modelo de Redes

Uma das características da maior parte dos SIEM é a capacidade que estes têm de lidar com a arquitectura de rede de determinada organização. Com esta capacidade, as empresas podem de alguma forma importar o seu modelo de rede na plataforma de forma a enriquecer a correlação. Normalmente este modelo de redes permite a definição de zonas de rede, blocos de IPs e os

próprios IPs das máquinas. Assim é mais fácil perceber todo o tráfego interno e externo que está a passar pela rede.

O resultado de um *scan* à rede, por uma ferramenta de *scan* de vulnerabilidades permite retirar muita informação que pode ser importada nas plataformas de SIEM. No entanto, quando o modelo de redes não está especificamente descrito e actualizado num ficheiro, o mais aconselhável é definir zonas e máquinas críticas de forma a facilmente se ter um enriquecimento das regras de correlação geradas pelas ferramentas de SIEM.

Categorizar os Activos e os Utilizadores

Tal como a aplicação de um modelo de redes, as ferramentas SIEM têm também a grande vantagem de se poder fazer uma categorização dos activos da infra-estrutura, assim como uma categorização dos utilizadores da mesma. Esta classificação permite que, aquando da criação de regras de correlação, possa ser tido em conta o impacto que estes activos e estes utilizadores têm para o sistema, por exemplo, as acções de utilizador (com classificação de impacto alta) sobre um activo (também com classificação de impacto alta) deve ser analisado sempre que se verifique uma alteração dos padrões normais.

Segundo a norma ISO 27001, um activo é qualquer recurso que representa valor para a organização. Estes activos devem ser classificados tendo em conta os princípios da integridade, disponibilidade e confidencialidade e podem ser considerados aspectos como: severidade do activo; susceptibilidade do activo perante ataques específicos; historial dos antigos ataques; e criticidade com base na importância que o activo tem na infra-estrutura.

Já os utilizadores devem ser classificados consoante o seu papel na organização e os privilégios que esse papel implica. É comum ter em conta alguns administradores com tarefas e responsabilidades mais críticas.

Operação e Administração

Configurar Regras de Gestão Interna da Plataforma

Algumas das preocupações que as empresas devem ter em relação aos seus sistemas estão relacionadas com a sua gestão e manutenção. Todos os sistemas, incluindo os SIEM, estão assentes sobre *hardware* e a qualquer altura alguma componente pode falhar. É importante que sejam criadas regras que permitam monitorizar esse *hardware*. O RAID de discos é um dos exemplos pois dependendo do tipo de RAID, quando um dos discos falha o sistema continua em

funcionamento mas é importante que o disco danificado seja rapidamente substituído para que sejam evitados danos maiores. Outro exemplo é as ventoinhas de refrigeração que, apesar do comum arrefecimento nos centros de dados, permitem que as temperaturas dos processadores não atinjam limites máximos que possam levar à avaria do mesmo.

Existem várias formas de fazer este tipo de monitorização. Por norma, as *appliances* contêm uma interface de rede que permite o acesso a uma consola de gestão. Entre outras coisas, estas consolas permitem fazer a gestão de todo o *hardware* permitindo assim a configuração ou definição de mecanismos de monitorização do mesmo.

Configurar Conteúdos

Estando reunidas todas as condições passa-se então para a configuração de algum conteúdo para que se possa tirar partido de todo o trabalho de implementação das fases anteriores. As ferramentas de SIEM contêm mecanismos poderosos de elaboração de relatórios e em complemento a este recurso contêm *dashboards*, que são conjuntos de gráficos que permitem a visualização mais amigável de determinadas situações que estão a ocorrer na plataforma. Os relatórios e *dashboards* são elaborados pegando na base dos casos de uso anteriormente definidos no planeamento do projecto. A definição destes relatórios e *dashboards* permite, entre outras coisas, garantir que os objectivos do projecto foram atingidos e se alguma componente do processo de implementação falhou. Por exemplo, se nos casos de uso se identificou que no final do projecto seria necessário verificar todas as máquinas infectadas com vírus e não se verificam eventos que permitam analisar estas situações, então é necessário voltar a fazer uma análise aos eventos de antivírus e identificar qual o problema.

Definir Regras de Correlação

As regras de correlação, tal como o nome indica, são regras que contêm condições definidas e correlacionam eventos à medida que estes são processados na ferramenta SIEM.

Sendo a correlação uma das componentes dos SIEM, é normal que esta tarefa de definição de regras de correlação seja uma tarefa morosa. São estas as regras que vão despoletar alguma acção sempre que um comportamento definido nas condições for identificado. As ferramentas de SIEM contêm regras de correlação já definidas que podem e devem ser utilizadas. No entanto estas regras, habitualmente baseadas em heurísticas, devem ser analisadas com algum rigor pois as suas condições não se ajustam a todos os tipos de negócio.

Uma regra de correlação de utilização comum é a de detecção de ataques de *Brute-Force.* Esta regra é despoletada quando um determinado número de tentativas falhadas de login for

identificado num determinado intervalo de tempo. É claro que para a identificação destes ataques não era necessário um SIEM. No entanto, uma vez que temos eventos de vários sistemas, é possível correlacionar as tentativas de *Brute-Force* em sistemas distintos e identificar tentativas de *login* por *Brute-Force* que sem correlação de eventos não seria possível.

Analisar e Corrigir os Falsos Positivos e Falsos Negativos

Um dos grandes desafios a superar na área da segurança é optimização dos alertas que são gerados, tal como foi referido inicialmente. A má configuração de um equipamento pode levar a um alerta de ataque quando na verdade o que está a acontecer na rede não é um ataque. Estes falsos alarmes têm de ser analisados e corrigidos ao longo do tempo para que se possam ajustar as regras de correlação, para que estas apenas sejam despoletadas quando verdadeiros ataques estejam a acontecer. Normalmente as regras de correlação que mais falsos positivos geram são as baseadas em *firewalls* e IDSs devido à sua função na rede. Portanto é necessário bastante cuidado quando se configuram regras baseadas nestas tecnologias, pois mais tarde pode ter de se perder muito tempo com a verificação dos falsos positivos.

Ao contrário dos falsos positivos, os falsos negativos são mais difíceis de detectar sem que o seu efeito malicioso tenha acontecido. Depois de definida uma regra de correlação, é conveniente que se façam vários testes; caso contrário um ataque pode vir a ser bem sucedido e só depois de identificado é que se verifica que estamos perante um falso negativo.

Definir Política de Cópias de Segurança

Tal como em todos os sistemas, uma das principais preocupações a ter nas implementações de SIEM é a definição de uma política de *backup* de configuração.

A qualquer momento se pode perder uma configuração, seja por debilidades do *hardware*, do *software*, por distracção do utilizador, ou por qualquer outro motivo. A realidade é que se houver uma política de *backup*, a reposição das configurações torna-se na maioria dos casos uma tarefa fácil. Caso contrário existe a necessidade de se proceder à reconfiguração dos sistemas.

Capítulo V – Validação do Método de Implementação de SIEMs

A definição de parâmetros de validação do método de implementação de SIEMs proposto anteriormente é uma das actividades prevista pela metodologia de investigação utilizada. Dadas as limitações de tempo inerentes a este trabalho de dissertação, a validação do método proposto não pôde ser concluída da melhor forma. No entanto, a principal forma de validação será através da implementação de uma plataforma de SIEM baseada em tecnologia ArcSight.

Implementação de um SIEM numa empresa Portuguesa

O projecto de implementação que vai permitir ao aluno fazer uma primeira validação do método de implementação que propôs já se encontra em curso numa grande empresa Portuguesa dispersa por quatro locais diferentes do país (3 distritos a norte do rio Tejo) sendo este conduzido pela Unisys Portugal. A fase de planeamento identificada no método de implementação de SIEMs anteriormente descrito, encontra-se terminada e é seguidamente detalhada.

Esta implementação teve como ponto de partida uma plataforma de SIEM anterior. No entanto, os problemas de implementação que nunca puderam ser resolvidos levaram a que a plataforma nunca operasse conforme as expectativas sendo que a mudança de tecnologia para a tecnologia ArcSight foi inevitável.

O principal objectivo da organização passava pela implementação de uma plataforma SIEM com capacidade para: consolidação e arquivo de eventos, correlação de eventos, análise em tempo real, retenção de evidências *online* durante 12 meses e geração de relatórios de conformidade de acordo com a norma PCI DSS.

No âmbito deste projecto a organização identificou que todo o seu negócio deveria ser monitorizado, desde o perímetro da sua rede à rede interna, incluindo aplicações e colaboradores.

Relativamente aos *use cases*, para além dos que foram mencionados no capitulo IV foram também definidos outros *use cases* relacionados com as aplicações proprietárias da organização e com a monitorização de contas de colaboradores com privilégios máximos como, por exemplo, administradores de sistemas e administradores de bases de dados.

Para a definição dos tipos de eventos a monitorizar foi decidido que seriam monitorizados todos os eventos aplicacionais assim como os eventos dos sistemas operativos que suportam essas aplicações, sejam elas proprietárias ou não. Com base nesta politica de monitorização foram

identificados aproximadamente 100 equipamentos que passam por sistemas operativos Windows e Linux, *firewalls* internas e de perímetro, bases de dados, anti-vírus, detectores de intrusões, aplicações proprietárias, serviço de DHCP, VPNs e *mainframe.*

Após identificados os equipamentos e o tipo de eventos a monitorizar foram feitas estimativas aproximadas para o número de eventos gerados por segundo tendo-se chegado a um valor de 750 com picos de eventos a poder atingir os 1500.

A definição do período de retenção dos eventos está directamente relacionada com as normas que a organização tem de cumprir. No entanto definiu-se um limite mínimo superior ao exigido dado que o total do número de eventos gerados não era exagerado. O período de retenção de informação foi assim fixado em 12 meses.

Foi definida uma arquitectura física de acordo com a figura 19. Verifica-se que esta solução é baseada maioritariamente em *appliances* da ArcSight e que estão a ser monitorizados 4 locais distintos, sendo que os locais 2, 3 e 4 (ver figura 19) possuem um servidor que suporta os conectores de cada local. Estes conectores processam os eventos e têm a capacidade de os enviar paralelamente para o componente de gestão de eventos denominado Logger, para a componente de correlação denominada Express e existe também a possibilidade de enviar os eventos para designado "Ambiente de Qualidade" em que podem ser feitos inúmeros testes sem que os sistemas em produção sejam afectados. Tal como acontece nos locais 2, 3 e 4, o mesmo acontece no local 1.Verificamos também que ambos os componentes (gestão e correlação) possuem consolas independentes que permitem que os analistas executem o trabalho consoante as suas responsabilidades. Na parte superior esquerda da imagem verifica-se um rectângulo com a representação de alguns serviços utilizados pela ferramenta, de forma a melhorar o seu desempenho e enriquecer alguma da informação que é possível consultar.

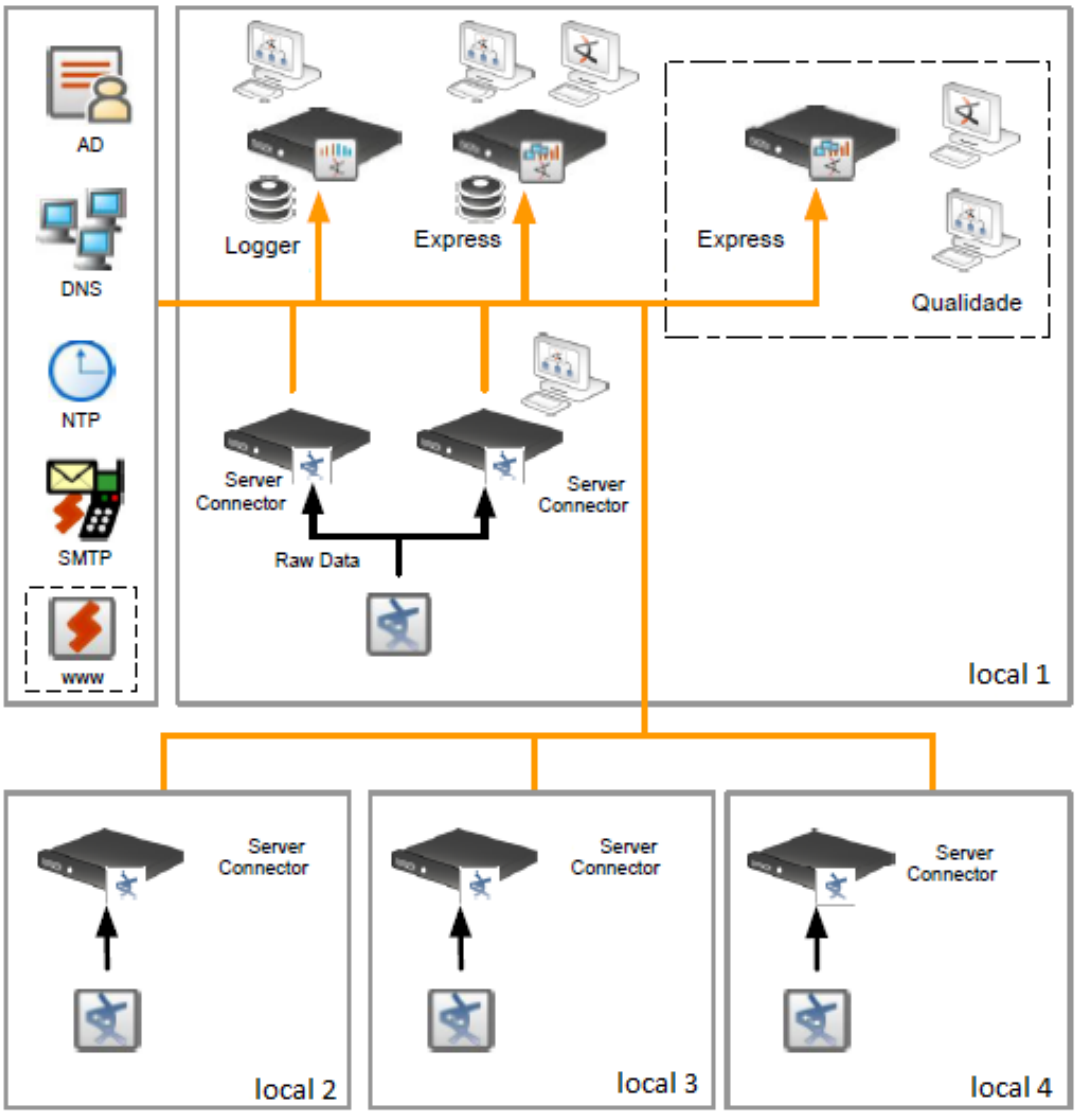


Figura 19: Arquitectura a Implementar

A definição de perfis e utilizadores resultou em dois tipos de perfis distintos: o perfil de administrador e o perfil de operador. Para além disso foi ainda identificado um perfil de auditor que se verificou poder ser definido em pormenor ao longo do projecto. Para o perfil de administrador foram identificados dois utilizadores e para o perfil de operadores foram identificados 3 utilizadores. Os administradores têm principalmente a responsabilidade de garantir a sustentabilidade da plataforma, garantindo o seu crescimento em termos de abrangência de equipamentos a monitorizar e para além disso têm a responsabilidade de manter a plataforma actualizada de forma garantir o bom funcionamento da mesma. Já os operadores têm a principal responsabilidade de desenvolver conteúdos, tal como regras ou relatórios, que lhes permitam fazer análise das ocorrências na infra-estrutura monitorizada.

Neste momento o projecto encontra-se na fase de instalação onde a recolha de informação está em curso, junto da equipa de projecto da organização em causa.

Recolha e análise de resultados

Uma vez que não foi possível recolher valores concretos para o a implementação a decorrer com base no método apresentado, foram recolhidos valores de um outro projecto de implementação de SIEM em que o aluno esteve também envolvido ao longo do período em que foi desenvolvida a dissertação.

Este projecto ocorreu numa das grandes organizações portuguesas de telecomunicações e o grande objectivo da mesma com aquisição de uma plataforma SIEM, é a centralização de eventos da sua infra-estrutura de forma a garantir que o seu centro de operações de segurança interno consegue facilmente identificar situações anómalas e inesperadas. Para além da monitorização da intra-estrutura interna, a empresa em causa vende também os seus serviços de monitorização para alguns dos seus clientes, ficando assim responsável pela identificação e neutralização de eventuais ataques.

A definição de *Use Cases* assim como a definição do período mínimo de retenção de eventos não aparentavam ser um problema na fase inicial do projecto. No entanto, com o decorrer do tempo e com o aumento das vendas dos serviços de monitorização de segurança, o número de eventos de entrada na plataforma ultrapassa os 4000 eventos por segundo (provenientes de cerca de 400 equipamentos), fazendo com que o desempenho da plataforma assim como o período de retenção de eventos não seja o mais adequado face às necessidades conhecidas.

A arquitectura definida para a plataforma de SIEM na organização em causa encontra-se representada na figura 20. Verifica-se que os eventos gerados nos diversos equipamentos são recolhidos através de conectores e encaminhados para o ArcSight Logger que tem como principal função a retenção dos eventos. Para além disso o ArcSight Logger envia os eventos previamente definidos importantes para a correlação, para o ArcSight ESM Manager que por sua vez os armazena numa base de dados representada por ArcSight Database. Verifica-se também uma componente designada ArcSight Connector Appliance que tem como função a gestão remota dos conectores. Pode-se constatar que toda a comunicação entre os equipamentos da ArcSight é feita através de SSL e todos os componentes possuem consolas para administração e consulta da informação.

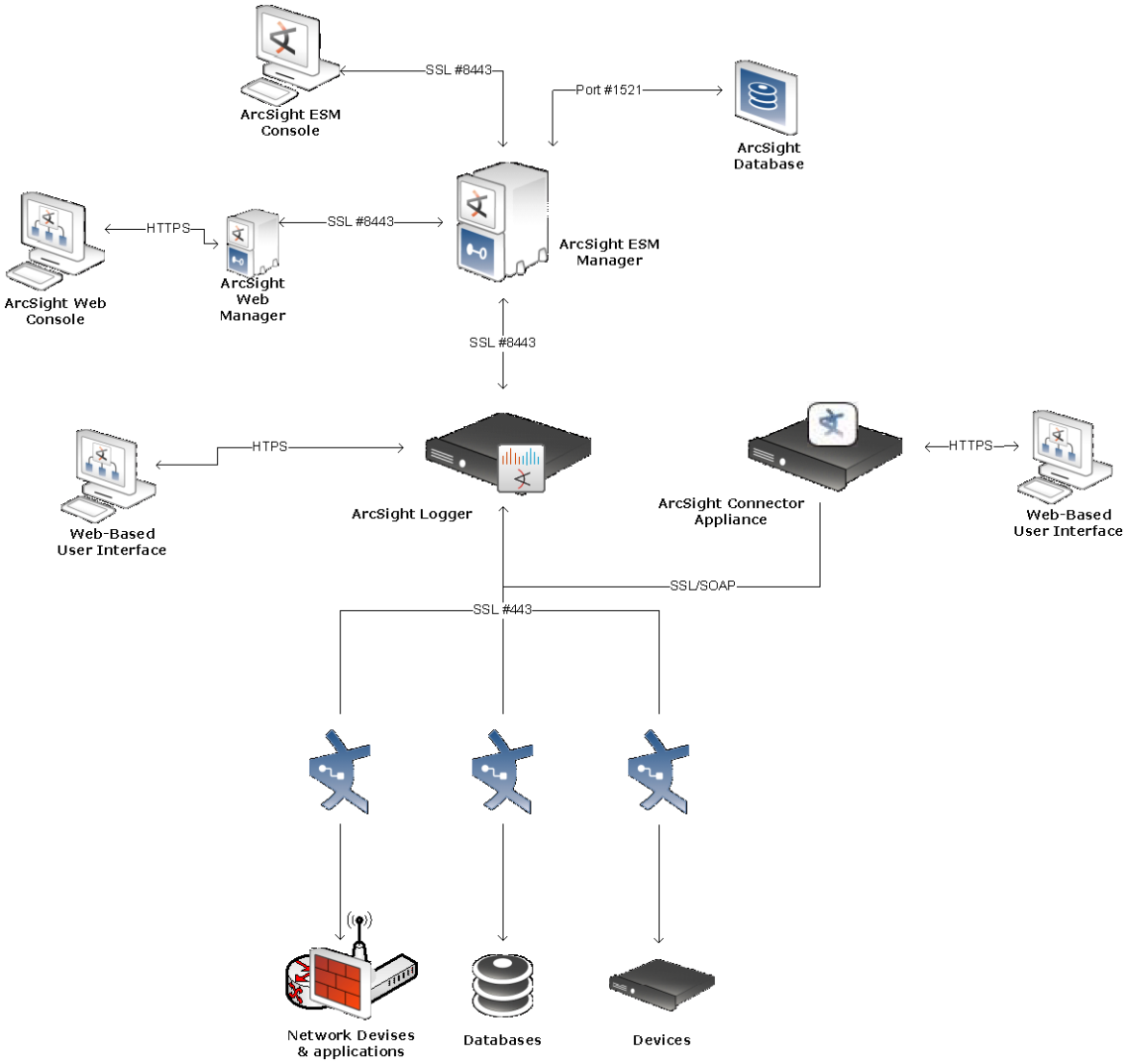


Figura 20: Arquitectura ArcSight em projecto de Referência

Neste projecto foram retiradas métricas de carácter analítico e métricas de carácter tecnico cujo objectivo é servirem como referência futura de forma a poder ser feita uma avaliação do funcionamento da plataforma actualmente a ser implementada com base no método apresentado.

Métricas de carácter analítico:

Na tabela 7 encontram as 10 portas mais utilizadas em comunicações na organização em causa assim como os protocolos associados a estas portas. Estes valores correspondem a um período de um mês e pode-se concluir que a porta 80 é a mais solicitada, de acordo com o tipo de eventos que a plataforma SIEM está a receber.

Porta	Protocolo	Número de Eventos (11-09-2011 a 11-10-2011)
80	TCP	2547314542
53	UDP	1241144463
514	UDP	786517683
161	UDP	378678611
443	TCP	282394182
8080	TCP	260636798
25	TCP	236955578
3478	UDP	181707610
445	TCP	169109429
110	TCP	154800124

Tabela 7: Top Portas Utilizadas

Na tabela 8 podem ser verificados os eventos mais recolhidos durante um mês na plataforma de SIEM. Para além disso podemos verificar qual o produto e a tecnologia que regista esses eventos. Analisando a tabela 8 pode-se concluir que os produtos que geram mais eventos são as *firewalls* e os *proxies* sendo que o evento designado "*accept"* foi o que se verificou mais vezes.

Nome do Evento	Tecnologia	Produto	Número de Eventos (11-09-2011 a 11-10-2011)
accept	Check Point	VPN-1 & FireWall-1	2836166236
URI Access Attempt	Optenet	Optenet Webfilter	1417902257
Deny udp	CISCO	PIX	782150434
drop	Check Point	VPN-1 & FireWall-1	536473846
Built inbound TCP connection	CISCO	FWSM	509629950
Built outbound UDP connection	CISCO	FWSM	472219725
Built outbound TCP connection	CISCO	FWSM	404900116
PATH	Unix	auditd	395436453
EOE	Unix	auditd	346707009
SYSCALL	Unix	auditd	346228878

Tabela 8: Top Eventos gerados

Métricas de carácter técnico:

Na tabela 9 podem-se verificar os 15 produtos e as tecnologias associadas que geraram mais eventos ao longo de um mês na organização em causa. Mais uma vez se verifica que as *firewalls* são as responsáveis por mais eventos dentro da plataforma SIEM, seguindo-se os sistemas operativos Unix e os servidores de *proxies.*

Tecnologia	Produto	Número de Eventos (11-09-2011 a 11-10-2011)
Check Point	VPN-1 & FireWall-1	3397810744
CISCO	FWSM	1784229906
Unix	auditd	1555760141
Optenet	Optenet Webfilter	1417902257
CISCO	PIX	802133434
Unix	Unix	195632088
CISCO	CiscoRouter	96972066
Fortinet	Fortigate	33729835
CISCO	ASA	22713514
Enterasys	Dragon	18664911
McAfee	ePolicy Orchestrator	12944483
Microsoft	Exchange	2981580
Check Point	SmartDefense	2673573
SmartSolution	SmartAgent	2308638
IP Flow	IP Flow	2065260

Tabela 9: Top produtos com mais eventos

Capítulo VI - Conclusões

Ao longo deste documento foram apresentados os sistemas de detecção e monitorização de intrusões, realçando a sua evolução e importância actual para as empresas. Pela revisão de literatura apresentada foi possível detectar a lacuna que este trabalho pretendeu colmatar: escassez de informação e orientações estruturadas para orientar o implementador destes sistemas, guiando a sua atenção para os aspectos que deve ter em atenção para assegurar que a empresa fica devidamente protegida de ataques internos e externos à informação que gere e aos seus sistemas.

Devido a restrições temporais inerentes à realização de um trabalho de dissertação, não foi possível validar o método proposto da melhor forma possível. No entanto, também neste documento foi apresentada a forma como a fase de Planeamento foi implementada num projecto que teve inicio em Outubro e para além disso são apresentadas métricas retiradas de um outro projecto de SIEM já implementado, para referência em validações futuras.

Depois de validado, o método deverá ser aplicado em estudos posteriores para afinação das actividades propostas a contextos diversos incluindo diferentes (i) dimensões da empresa, (ii) cultura empresarial, (iii) dimensão dos sistemas da empresa, (iv) modelo de negócio, entre outros contextos.

O trabalho realizado tem como principais resultados, a apresentação de informação detalhada sobre os sistemas de detecção de intrusões e de um método para orientação das actividades de implementação de sistemas de SIEM

A informação detalhada sobre os sistemas de detecção de intrusões pode tornar-se uma fonte de informação muito útil para quem quiser compreender a utilidade e funcionamento destes sistemas.

Depois de validado e afinado, o método será particularmente útil para os profissionais que pretendam entender o processo de implementação de um sistema de detecção de intrusões bem como obter recomendações práticas para as actividades que necessitará executar de forma a assegurar a implementação adequada destes sistemas.

Referências Bibliográficas

Amorosi, D. (2011). Data Breach Spring. *Infosecurity*, *8*(3), 6-9. Elsevier Ltd. doi:10.1016/S1754-4548(11)70032-8

Andress, A. (2004). Intrusion Detection. *Surviving Security: How to Integrate People, Process and Technology* (Second Edi., p. 500). AUERBACH.

Angelino, R. (n.d.). *Using events-per-second as a factor in selecting SEM tools*. Retrieved from http://www.infosecwriters.com/text_resources/pdf/events_per_second.pdf

ArcSight. (2007). Using Advanced Event Correlation to Improve Enterprise Security , Compliance and Business Posture. *Business*.

Axelsson, S. (2000). Intrusion Detection Systems : A Survey and Taxonomy. Department of Computer- Engineering, Chalmers University.

Bace, R. G. (2000). *Intrusion Detection* (p. 340). Macmillan Tachnical Publishing.

Barber, R., & Mell, P. (2001, June 1). Intrusion Detection Systems. *Computer Fraud*. doi:10.1016/S1361-3723(01)00614-5

Belcourt, M. (2006). Outsourcing — The benefits and the risks. *Most*, *16*, 269 - 279. doi:10.1016/j.hrmr.2006.03.011

Binde, B. E., Mcree, R., & Connor, T. J. O. (2011). *Assessing Outbound Traffic to Uncover Advanced Persistent Threat* (p. 35). Retrieved from http://www.sans.edu/student-files/projects/JWP-Binde-McRee-OConnor.pdf

Butler, J. M. (2009). Benchmarking Security Information Event Management (SIEM). *Event (London)*. Retrieved from http://www.sans.org/reading_room/analysts_program/eventMgt_Feb09.pdf

Cavusoglu, H., & Zhang, J. (2008). Security Patch Management: Share the Burden or Share the Damage? *Management Science*, *54*(4), 657-670. doi:10.1287/mnsc.1070.0794

Choudhary, V. (2007). Software as a Service : Implications for Investment in Software Development The Paul Merage School of Business. *Sciences-New York*, 1-10.

Davi, L., Dmitrienko, A., Sadeghi, A.-reza, & Winandy, M. (2011). Privilege Escalation Attacks on Android. *System Security Lab*.

Denning, D. E. (1987). An Intrusion-Detection Model. *IEEE Transactions on Software Engineering*, *SE-13*(2), 222-232. doi:10.1109/TSE.1987.232894

Gabriel, R., Hoppe, T., Pastwa, A., & Sowa, S. (2009). Analyzing Malware Log Data to Support Security Information and Event Management: Some Research Results. *2009 First International Confernce on Advances in Databases, Knowledge, and Data Applications*, 108-113. Ieee. doi:10.1109/DBKDA.2009.26

Hofmeyr, S. A., Forrest, S., & Somayaji, A. (1998). *Intrusion Detection using Sequences of System Calls* (p. 25).

Hong, L. (2009). Immune Mechanism Based Intrusion Detection Systems. *2009 International Conference on Networks Security, Wireless Communications and Trusted Computing* (pp. 568-571). Ieee. doi:10.1109/NSWCTC.2009.22

Huang, M.-Y., Jasper, R. J., & Wicks, T. M. (1999). A large scale distributed intrusion detection framework based on attack strategy analysis. *Computer Networks*, *31*(23-24), 2465-2475. doi:10.1016/S1389-1286(99)00114-0

Järvinen, P. (2007). Action Research is Similar to Design Science. *Quality Quantity*, *41*(1), 37-54. Springer. doi:10.1007/s11135-005-5427-1

Kent, K., & Souppaya, M. (2006). Guide to Computer Security Log Management. *Nist Special Publication*.

Kim Zetter. (2011). No Title. Retrieved September 28, 2011, from http://www.wired.com/threatlevel/2010/03/source-code-hacks/

Mathew, D. (2002). Choosing an Intrusion Detection System that Best Suits your Organization. *SANS Institute*.

McAfeeLabs. (2011). *Relatório da McAfee sobre Ameaças : Segundo trimestre de 2011* (p. 24).

Miguel, A. (2006). *Gestão Moderna de Projectos* (2nd ed., p. 440). FCA.

Miller, D. R., Harris, S., Harper, A. A., VanDyke, S., & Blask, C. (2011). *Security Information and Event Management (SIEM) Implementation*. *Information Security* (p. 465). McGraw - Hill Companies.

Myers, J., Grimaila, M., & Mills, R. (2011). Log-Based Distributed Security Event Detection Using Simple Event Correlator. *Proceedings of the 44th Hawaii International Conference on System Sciences* (pp. 1-7).

Nakamura, E. T., & Geus, P. L. D. (2007). *Segurança de Redes em Ambientes Cooperativos* (p. 488). Novatec.

Nicolett, M. (2010). *Critical Capabilities for Security Information and Event Management Technology* (p. 16).

Payment Card Industry Data Security Standard. (2009, April). *Card Technology Today*. doi:10.1016/S0965-2590(09)70094-5

Plohmann, D., Padilla, E. G., & Leder, F. (2011). Botnets: Detection, Measurement, Disinfection & Defence. *Information Security*.

PunditNetworks. (2011). Opplev neste generasjons SIEM løsning. Retrieved September 12, 2011, from http://www2.pundit.no/wp-content/uploads/2011/03/logg-siem2.gif

Ragsdale, D. J., Carver, C. A., Humphries, J. W., & Pooch, U. W. (1999). Adaptation Techniques for Intrusion Detection and Intrusion Response Systems. *Methodology*.

Rashid, F. Y. (2011). 10 Biggest Data Breaches of 2011 So Far. Retrieved from http://www.eweek.com/c/a/Security/10-Biggest-Data-Breaches-of-2011-So-Far-175567/

Simão, A. M. D. L., Sícoli, F. C., Melo, L. P. D., Deus, F. E. D., Timóteo, R., & Júnior, D. S. (2011). Aquisição de Evidências Digitais em Smartphones Android, *2*(1), 92-99.

Swift, D. (2006). A Practical Application of SIM/SEM/SIEM Automating Threat Identification. *Information Security*.

Tankard, C. (2011). Advanced Persistent threats and how to monitor and deter them. *Network Security*, (8), 16-19. Elsevier Ltd. doi:10.1016/S1353-4858(11)70086-1

Tyler, G., & Wu, T. M. (2009). Intrusion Detection Systems. *IATAC*, 93. doi:10.1016/j.istr.2005.08.001